Explicit Stationarity Conditions and Solution Characterization for Equilibrium Problems with Equilibrium Constraints

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

doctor rerum naturalium
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Herrn Thomas Michael Surowiec
geboren am 09.07.1982 in Passaic, New Jersey, USA

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Peter Frensch

Gutachter:
1. PD Dr. René Henrion
2. Prof. Dr. Werner Römisch
3. Prof. Dr. Jiří V. Outrata

eingereicht am: 12. August 2009
Tag der mündlichen Prüfung: 27. Januar 2010

To my mother and father,
without whom this would
never have been possible.

Acknowledgments

I would first and foremost like to thank my two academic advisors René Henrion and Jiří V. Outrata for their unwavering support of me and my project. Their patience and dedication, along with their deep knowledge of the theory of optimization and variational analysis, allowed me to rapidly immerse myself in what was a relatively new subject to me when I began researching three years ago.

In addition, I wish to thank Werner Römisch, whose weekly optimization seminars provided me with a semesterly forum for presenting my latest results and whose thought-provoking questions guided me along the way, as well as Boris S. Mordukhovich and Michal Červinka, with both of whom I had many fruitful discussions. Furthermore, I would like to thank Darinka Dentcheva for first showing me the beauty of optimization theory and Douglas I. Bauer for giving me my first glimpses into the world of mathematics.

This dissertation is the result of three years of work at the Humboldt University of Berlin within the international research training group RTG 1128 "Multiphase Problems" funded by the German Research Foundation (DFG), to whom I am indebted for the bestowment of finanical support. The research environment made possible under the auspices of the DFG was integral in my development during the last three years. I would personally like to thank all my colleagues for their helpful discussions, especially Kshitij Kulshreshta, whose insight led to the creation of Example 3.1. I also wish to thank Michael Hintermüller for providing me with a new position at Humboldt three months before the end of my contract with the research training group; his patience and understanding during this time allowed for the completion of this dissertation.

Lastly, it should not go without mentioning my gratefulness to my family (especially my brother Paul for proofreading the introduction) and friends, on both sides of the Atlantic, for all their love and support and giving me the strength to continue striving forwards.

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit Gleichgewichtsproblemen unter Gleichgewichtsrestriktionen, sogenannten EPECs (Englisch: Equilibrium Problems with Equilibrium Constraints). Konkret handelt es sich um gekoppelte Zwei-Ebenen-Optimierungsprobleme, bei denen Nash- Gleichgewichte für die Entscheidungen der oberen Ebene gesucht sind. Ein Ziel der Arbeit besteht in der Formulierung dualer Stationaritätsbedingungen zu solchen Problemen. Als Anwendung wird ein oligopolistisches Wettbewerbsmodell für Strommärkte betrachtet.

Zur Gewinnung qualitativer Hypothesen über die Struktur der betrachteten Modelle (z.B. Inaktivität bestimmter Marktteilnehmer) aber auch für mögliche numerische Zugänge ist es wesentlich, EPEC-Lösungen explizit bezüglich der Eingangsdaten des Problems zu formulieren. Der Weg dorthin erfordert eine Strukturanalyse der involvierten Optimierungsprobleme (constraint qualifications, Regularität), die Herleitung von Stabilitätsresultaten bestimmter mengenwertiger Abbildungen und die Nutzung von Transformationsformeln für die sogenannte Ko-Ableitung. Weitere Schwerpunkte befassen sich mit der Beziehung zwischen verschiedenen dualen Stationaritätstypen (S- und M-Stationarität) sowie mit stochastischen Erweiterungen der betrachteten Problemklasse, sogenannten SEPECs.

Abstract

This thesis is concerned with equilibrium problems with equilibrium constraints or EPECs. Concretely, we consider models composed by coupling together two-level optimization problems, the upper-level solutions to which are non-cooperative (Nash-Cournot) equilibria. One of the main goals of the thesis involves the formulation of dual stationarity conditions to EPECs. A model of oligopolistic competition for electricity markets is considered as an application.

In order to profit from qualitative hypotheses concerning the structure of the considered models, e.g., inactivity of certain market participants at equilibrium, as well as to provide conditions useful for numerical procedures, the ablilty to formulate EPEC solutions in relation to the input data of the problem is of considerable importance. The way to do this requires a structural analysis of the involved optimization problems, e.g., constraints qualifications, regularity; the derivation of stability results for certain multivalued mappings, and the usage of transformation formulae for so-called coderivatives. Further important topics address the relationship between various dual stationarity types, e.g., S- and M-stationarity, as well as the extension of the considered problem classes to a stochastic setting, i.e., stochastic EPECs or SEPECs.

Contents

Chapter 1

Introduction

The study of optimal decision making in settings with many players seeking to maximize their utilities within a hierarchical framework, extends back to the founding works of the theory of non-cooperative games. Indeed, in Cournot [1838], a setting in which two decision makers producing a homogeneous product wishing to maximize their profits and competing without cooperation was considered. Later von Stackelberg [1934] studied competition in a hierarchical setting between a stronger (*leader*) and a weaker (*follower*) decision maker. In Morgenstern and von Neumann [1944], the mathematical formalization of these and other games was considered and in the seminal work of Nash [1951], the concept of a non-cooperative or Nash-Cournot equilibrium was developed. More precisely, a non-cooperative or Nash-Cournot equilibrium is a profile of strategies for which each involved decision maker maximizes its utility function, which is dependent on the choices of the other decision makers, provided all other decisions makers also choose their optimal decisions from this profile (cf. Osborne and Rubinstein [1994, Definition 14.1]). It was later demonstrated by Harker [1984], that variational inequalities (cf. Facchinei and Pang [2003]) are ideal for modeling oligopolistic behavior; an example of a Nash-Cournot game. This led to the realization that a single-leader-multiple-follower game in which the followers are playing a Nash-Cournot game amongst themselves and are engaged in a Stackelberg game with the leader, can be aptly modeled by a so-called *mathematical program with equilibrium constraints* or MPEC (see Luo et al. [1997] and Outrata et al. [1998] and references therein). Outside the realm of game theory, we consider an MPEC to be any optimization problem in which an objective function is minimized or maximized over a feasible set defined, at least partially, via an equilibrium constraint, which is often a parametric variational inequality or generalized equation (cf. Robinson [1979]). Here, the parameters are the *upper-level* or *state* variables and the solutions to the equilibrium constraint are the *lower-level* or *control* variables. In this sense, an MPEC is a type of two-level optimization problem of an often highly non-convex nature, regardless of the input data.

Some decades after the revolutionary work of Nash, researchers began to apply the new developments in mathematical programming to model more complicated structures in which many leaders and followers were competing (see Sherali et al. [1982]). These games were called Stackelberg-Nash-Cournot and Multiple-Leader-Stackelberg games in Sherali et al. [1983] and Sherali [1984]. In the latter, the leaders and followers are engaged in a Stackelberg game and the leaders play a

Nash-Cournot game amongst themselves. However, in such models, the followers decisions are modeled by preprescribed "reaction curves" and not as solutions to parametric variational inequalities. Nevertheless, these early models can be considered to be the forerunners of what we now call *equilibrium problems with equilibrium constraints* or simply: EPECs and the new formulations allowed for the application and development of numerical methods for obtaining equilibria as well as the demonstration of existence results and an analysis of equilibria.

Following along the traditions outlined above, we define an EPEC to be a type of mathematical model composed by coupling together at least two MPECs through the equilibrium constraint. That is, each MPEC has the same equilibrium constraint and thus, the same lower-level variables. Moreover, the objective functions from each MPEC may also be perturbed by the upper-level variables of the other MPECs, thereby making EPECs amenable to classical Nash-Cournot settings. Following the parlance of game theory, an EPEC solution is then a type of equilibria in which the leaders and followers are playing a Nash-Cournot game amongst themselves and a Stackelberg game between each other.

In some situations, the solutions or stationary points of EPECs can be modeled by quasi-variational inequalities (cf. Pang and Fukushima [2005]) and in this sense solved numerically. Developments concerning numerical methods for EPECs can be found in Leyffer and Munson [April 2005] for a class of multiple-leader-single-follower EPECs and well as in Hu [2002], Ehrenmann [2004a], Su [2005] and Červinka [2008]. However, the methods proposed in these works, e.g., reformulation as a mixed-complementarity problem, non-linear program via relaxation, or homotopy methods, all require that the equilibrium constraint be modeled by a standard complementarity problem (cf. Facchinei and Pang [2003]), which in our setting will not be the case. One necessary component for the mentioned numerical methods, and one which we seek to provide in this thesis, involves the derivation of explicit dual stationarity conditions for EPECs in which the equilibrium constraint is more general than a standard complementarity problem.

More recently, researchers have become interested in the usage of multi-leader-follower games for modeling and understanding oligopolistic behavior in deregulated electricity spot markets, where the electricity generators represent the leaders and an independent system operator (ISO) players the role of a single follower. Such an application will serve as our prototype EPEC application and we draw our inspiration from two suggestions for modeling such behavior, namely the models developed and used in Hu [2002], Hu and Ralph [2005], and Hu et al. [2007] and those of Borenstein et al. [2000] and Escobar and Jofre [2005]. The models of Hu et al. take into consideration and analyze many of the intricacies of the market. However, none of them considers the effects of transmission losses on the outcomes. Such a consideration can drastically change the type of game considered if transmission losses are large enough, as was first noted in Borenstein et al. [2000] for a two-settlement, i.e., two-leader-single-follower, setting and later formalized to a more general setting by Escobar and Jofre [2005].

Another difficulty in the study of EPECs arises due to their highly non-convex nature. As noted in Hu and Ralph [2005], Hu et al. [2007], Ehrenmann [2004b], these models may possess many if not continua or "manifolds" (Ehrenmann

[2004b]) of equilibria. This presents a significant challenge to the analyst, as one would often wish to know as many solutions as possible. Accordingly, we seek to provide an analytical framework for the classification of general types of solutions to the spot market EPEC; a similar approach was developed in Henrion and Römisch [2007], yet we will consider a larger array of settings and a more complicated structure.

Therefore, in order to accomplish our goals of providing explicit multiplier -based stationarity conditions and utilizing these conditions to characterize solution types, we will need results pertaining to the structural properties of the considered EPEC, e.g., constraint qualifications and regularity of the involved feasible sets. In addition, knowledge of the stability (continuity) properties of certain multivalued mappings defined as the solution mappings to the equilibrium constraints, or as variants thereof, is of importance. Lastly, we will need *transformation formulae* for so-called coderivatives to normal cone mappings to not-necessarily polyhedral sets. These latter results represent the key to providing the explicit stationarity conditions.

Building on the initial results found in Henrion and Römisch [2007], an additional contribution of this thesis involves the development and discussion of stochastic EPECs or SEPECs. This is done in concordance with the growing desire of companies to plan more risk-aversely, as the inclusion of stochasticity allows one to make better predictions based on historical data. Therefore, we provide in Chapter 3 a formal definition of a SEPEC, define the stochastic spot market EPEC, thereby proving the well-definedness of the expected value function of the objectives, and at the end of Chapter 8, explicit stationarity conditions for a SEPEC in which the random parameters are assumed to be discretely distributed are derived.

The thesis is structured as follows. Chapter 2 provides the definitions, tools, and concepts from modern variational analysis and mathematical programming as well as our notational conventions. Afterwards, in Chapter 3, we provide the formal definitions of MPECs and EPECs. Chapter 3 also contains a discussion on the well-posedness of EPECs and our spot market EPEC in both deterministic and stochastic forms is introduced. Chapter 4 presents the two main types of stationarity conditions of solutions to MPECs and EPECs that we will be considering. Moreover, some extensions and equivalences between these stationarity concepts are demonstrated and the so-called Fréchet normal cone to the graph of the solution mapping to the equilibrium constraint is studied.

Motivated by the need for stability properties of solution mappings to equilibrium constraints for the application of stationarity conditions, Chapter 5 contains a few results on the stability of a related mapping called the perturbation mapping. In Chapter 6, we present the major results on transformation formulae for coderivatives to normal cone mappings from the recently published text Henrion et al. [2009c], which the author of this thesis co-wrote, along with some needed extensions and simplifications not contained within that text. As noted above, these results are integral in making the stationarity conditions outlined in Chapter 4 explicit. Chapter 7 represents a return to our spot market EPEC and is composed of results concerning its structural properties, i.e., results demonstrating the constraint qualifications and stability properties needed for the application of the stationarity conditions are presented. Finally, in Chapter 8, the results of all previous chapters are applied to the spot market EPEC. This includes explicit stationarity conditions for both the deterministic and stochastic spot market EPEC and a variety of examples for so-called two-settlement electricity spot markets. The examples demonstrate how the stationarity conditions can sometimes be reduced to simple constraint systems leading to the derivation of quantifying relationships between decision variables, thus allowing for an analytical classification of solutions via stationarity conditions. Comparisons between the stationarity concepts in the form of selectivity versus applicability are also provided.

Chapter 2

Preliminaries and Notation

2.1 Notation

We begin with some preliminary notation. Let $\mathbb{R}^s$ denote s-dimensional Euclidean space and $\mathbb{R}^s_+$ and $\mathbb{R}^s_-$ its non-negative and non-positive orthants, respectively. We will use $||\cdot||$ to represent the standard Euclidean norm, though when needed we may use an equivalent norm for certain arguments, and $\langle \cdot, \cdot \rangle$ the scalar product. For a vector $x \in \mathbb{R}^s$, we denote its components by x_i and given two vectors $x, y \in \mathbb{R}^s$, the notation "$0 \leq x \perp y \geq 0$" will sometimes be used to denote the complementarity relations: $x_i \geq 0$, $y_i \geq 0$ and $x_i y_i = 0$ for all $i = 1, \ldots, s$. The space of all $(s \times t)$-matrices with real-valued entries is denoted by $\mathbb{R}^{s \times t}$ and for $\mathcal{A} \in \mathbb{R}^{s \times t}$ and $I \subseteq \{1, \ldots, s\}$, $\mathcal{A}_I$ is the submatrix whose row-indices are in I. For a continuously differentiable function $f : \mathbb{R}^s \to \mathbb{R}$, we let $\nabla f(x) \in \mathbb{R}^{s \times 1}$ denote its gradient, however, if it is clear in context, we will avoid writing $\nabla f(x)^T d$ or $\langle \nabla f(x), d \rangle$ to denote the scalar product with some vector $d \in \mathbb{R}^s$, and simply write $\nabla f(x) d$. Given a continuously differentiable mapping $F : \mathbb{R}^s \to \mathbb{R}^t$, with components $F_i(x)$ and Jacobian $\nabla F(x) \in \mathbb{R}^{t \times s}$, we denote its transpose by $\nabla^T F(x)$ and its partial derivative with respect to x_i by $\nabla_{x_i} F(x)$. Note that $\nabla^T F_I(x)$ for some $I \subset \{1, \ldots, t\}$ means the transpose of $\nabla F_I(x)$. Given a *multifunction* $\Phi : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$, i.e., any function mapping $\mathbb{R}^s$ into the set of all subsets of $\mathbb{R}^t$, its *graph* is defined $\operatorname{gph} \Phi := \{(x, z) \in \mathbb{R}^s \times \mathbb{R}^t \,|\, z \in \Phi(x)\}$. For F and Φ as just defined, we will refer to any relation of the form: $0 \in F(x) + \Phi(x)$ as a *generalized equation* (see e.g., Robinson [1979]). Finally, the negative polar (dual) of a cone $K \subseteq \mathbb{R}^s$ is denoted $K^- := \{h^* \in \mathbb{R}^s \,|\, \langle h^*, h \rangle \leq 0, \ \forall h \in K\}$.

2.2 Basic Variational Geometry

We begin by defining the notions of variational geometry important for our study. For a closed set $C \subseteq \mathbb{R}^t$ and point $\bar{z} \in C$, we define the *contingent* or *Bouligand cone*

$$T_C(\bar{z}) := \operatorname*{Lim\,sup}_{\tau \searrow 0} \frac{C - \bar{z}}{\tau} = \left\{ d \in \mathbb{R}^t \,|\, \exists \tau_k \searrow 0, \exists d_k \to d : \forall k, \bar{z} + \tau_k d_k \in C \right\}.$$

Here, 'Lim sup' represents the Painlevé-Kuratowski upper/outer limit (cf. Rockafellar and Wets [1998]). In the event C is convex, this reduces to the standard tangent cone from convex analysis. Using the contingent cone, we define the *Fréchet normal cone*

$$\widehat{N}_C(\bar{z}) := [T_C(\bar{z})]^- .$$

By taking the outer limit of Fréchet normal cones in the following way

$$N_C(\bar{z}) := \operatorname*{Lim\,sup}_{\substack{z \to \bar{z} \\ z \in C}} \widehat{N}_C(\bar{z}),$$

we define the *Mordukhovich* or *limiting normal cone.* Note that if C is convex, then $\widehat{N}_C = N_C$ and the cones coincide with the classical notion of normal cone from convex analysis. We refer the reader to Aubin and Frankowska [1990] for many results on contingent cones and Mordukhovich [2006a] and Rockafellar and Wets [1998] for various results pertaining to all three objects.

2.3 Coderivatives of Multifunctions

A central part of our analysis relies on our ability to explicitely calculate certain generalized derivatives known as *coderivatives.* Given a multifunction $\Phi : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$, we define the *Mordukhovich coderivative* of Φ at $(\bar{x}, \bar{z}) \in \operatorname{gph} \Phi$ in (dual) direction $z^* \in \mathbb{R}^t$

$$D^*\Phi(\bar{x}, \bar{z})(z^*) := \{x^* \in \mathbb{R}^s \,|(x^*, -z^*) \in N_{\operatorname{gph} \Phi}(\bar{x}, \bar{z})\} .$$

Mordukhovich coderivatives, due to their robustness and full calculus, have been extensively studied given their usefullness in characterizing and checking the presence of certain stability properties of multifunctions. The large majority of these results can be found in Mordukhovich [2006a].

As we will need it for discussion, we also provide the definition of the *Fréchet coderivative* of Φ, denoted $\widehat{D}^*\Phi$, which states

$$\widehat{D}^*\Phi(\bar{x}, \bar{z})(z^*) := \left\{x^* \in \mathbb{R}^s \,\middle|(x^*, -z^*) \in \widehat{N}_{\operatorname{gph} \Phi}(\bar{x}, \bar{z})\right\} .$$

Given the Fréchet coderivatives, one can provide an equivalent definition of the Mordukhovich coderivative as follows (cf. Mordukhovich [2006a, Definition 1.32]: Let $(\bar{x}, \bar{z}) \in \operatorname{gph} \Phi$ and $\bar{v}^* \in \mathbb{R}^s$, then

$$D^*\Phi(\bar{x}, \bar{z})(\bar{v}^*) = \operatorname*{Lim\,sup}_{\substack{(z,v,v^*) \to (\bar{z},\bar{v},\bar{v}^*) \\ (x,z) \in \operatorname{gph} \Phi}} \hat{D}^*\Phi\,(x, z)\,(v^*),$$

Of course the various types of normal cones available in the literature allow for the derivation of a number of different coderivatives. However, in this thesis, we almost exclusively, with the exception of Chapter 6 Section 6.3, use the Mordukhovich coderivative and thus, we simply write “coderivative” for the Mordukhovich coderivative for the sake of brevity.

2.4 Stability Notions for Multifunctions

The presence of certain stability (continuity) properties of multifunctions occupies a central part of our study. Though there are many types of stability concepts for multifunctions, we are mainly concerned with the following fundamental notions. We begin with three Lipschitz-like continuity properties for general multifunctions. Let $\Phi : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$ be a multifunction and $(\bar{x}, \bar{z}) \in \operatorname{gph} \Phi$. Then we say that Φ has the *Aubin property*[1] at $(\bar{x}, \bar{z})$, if there exists neighborhoods $\mathcal{U}$ of $\bar{x}$ and $\mathcal{V}$ of $\bar{z}$ and a constant $\kappa > 0$ such that the following relation holds

$$d(z, \Phi(x')) \leq \kappa ||x - x'||, \forall z \in \mathcal{V} \cap \Phi(x), \forall x, x' \in \mathcal{U}.$$

Thus near $(\bar{x}, \bar{z})$, Φ enjoys a Lipschitz-like continuity property without being necessarily single-valued. Note that the last statement can be equivalently written in the following form

$$\Phi(x) \cap \mathcal{V} \subset \Phi(x') + L||x - x'||\mathbb{B} \quad \forall x, x' \in \mathcal{U}.$$

Next, if we fix x' in the previous definition by setting it equal to $\bar{x}$, we say that Φ is *calm* at $(\bar{x}, \bar{z})$; appropriately, this property is referred to as *calmness*. Note how calmness is a pointwise property, whereas the Aubin property is locally based. For multifunctions $Z : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^p$ defined

$$Z\,(\alpha, \beta) := \{x \in \mathbb{R}^p | G_1(x) = \alpha,\ G_2(x) \leq \beta\},$$

where $G_1 : \mathbb{R}^p \to \mathbb{R}^n$ and $G_2 : \mathbb{R}^p \to \mathbb{R}^m$ are continuous mappings, it is easy to see that calmness of Z at $(0, 0, \bar{x})$ for some $\bar{x}$ satisfying $G_1\,(\bar{x}) = 0$ and $G_2\,(\bar{x}) = 0$ is equivalent with the existence of $L, \varepsilon > 0$, such that

$$d(x, Z(0,0)) \leq L \left(\sum\nolimits_i |G_{1i}(x)| + \sum\nolimits_i \left[G_{2i}(x)\right]_+ \right) \quad \forall x \in \mathbb{B}\,(\bar{x}, \varepsilon)\ . \tag{2.1}$$

Here, $[y]_+ := \max\{y, 0\}$. Finally, if $\mathcal{V} \equiv \mathbb{R}^t$ in this definition of calmness, we say that Φ is *upper-Lipschitzian* at $(\bar{x}, \bar{z})$.

As we will study the stability properties of solution mappings to generalized equations, we will also need the following stability property. Begin by defining the solution mapping

$$S(x) := \left\{ z \in \mathbb{R}^t \,| 0 \in F(x, z) + N_C(z) \right\},$$

where $F : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}^t$ is at least continuously differentiable and C is some closed convex set. Let $(\bar{x}, \bar{z}) \in \operatorname{gph} S$ and define the multifunction $\Sigma : \mathbb{R}^t \rightrightarrows \mathbb{R}^t$ via a local partial linearization of the generalized equation corresponding to S, i.e.,

$$\Sigma(\xi) := \left\{ z \in \mathbb{R}^t \,| \xi \in F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) + N_C(z) \right\}.$$

[1] also called *pseudo-Lipschitzness* and the *Lipschitz-like* property, see Aubin and Frankowska [1990], Klatte and Kummer [2002b] and Mordukhovich [2006a], resp.

If there exist neighborhoods $\mathcal{W}$ of $0 \in \mathbb{R}^t$ and $\mathcal{V}$ of $\bar{z}$ such that the map $\xi \mapsto \Sigma(\xi) \cap \mathcal{V}$ is *single-valued and Lipschitz continuous* on $\mathcal{W}$ with modulus κ, then the generalized equation is is called *strongly regular* at $(\bar{x}, \bar{z})$.

The concept of strong regularity is due to [Robinson, 1980] and has the implication that for any $\varepsilon > 0$ there exist neighborhoods $\mathcal{U}_\varepsilon$ of $\bar{x}$ and $\mathcal{V}_\varepsilon$ of $\bar{z}$ such that the mapping $x \mapsto \sigma(x) := S(x) \cap \mathcal{V}_\varepsilon$ is single-valued and Lipschitz on $\mathcal{U}_\varepsilon$ with Lipschitz modulus $(\kappa + \varepsilon)L$, where L is the uniform Lipschitz modulus of $F(\cdot, z)$ on $\mathcal{U}_\varepsilon$ for all $z \in \mathcal{V}_\varepsilon$ (see [Robinson, 1980], Theorem 2.1).

2.5 Constraint Qualifications

For a set $C := \{z \in \mathbb{R}^t \,|A_j(z) \leq 0, j = 1, \ldots, p\}$, where $A : \mathbb{R}^t \to \mathbb{R}^p$ is continuously differentiable. Let $\hat{z} \in C$ and define

$$\begin{aligned} I(\hat{z}) &:= \{j | A_j(\hat{z}) = 0\} \quad \text{“active index set”} \\ L(\hat{z}) &:= \{1, \ldots, p\} \setminus I(\hat{z}) \quad \text{“inactive index set”} \end{aligned}$$

Then we say that the *Linear Independence Constraint Qualification* (LICQ) holds at $\bar{z} \in C$ if $\nabla A_{I(\bar{z})}(\bar{z})$ is a surjective mapping. If the collection of gradients $\{\nabla A_j(\bar{z})\}_{j \in I(\bar{z})}$ is only positively linearly independent, then we say that the *Mangasarian -Fromowitz Constraint Qualification* (MFCQ) holds at $\bar{z} \in C$. The *Constant Rank Constraint Qualification* (CRCQ) is said to hold at $\hat{z}$ provided there exists a subset $\mathcal{V} \subset C$ of $\hat{z}$ such that for each $K \subset I(\hat{z})$, $\operatorname{rank}\{\nabla A_K(z)\}$ remains constant on $\mathcal{V} \subset C$. It is important to note that

$$LICQ \Rightarrow MFCQ, CRCQ \quad MFCQ \not\Leftrightarrow CRCQ$$

The CRCQ was initially provided in Janin [1984] and has recently been investigated in Lu [2009].

For C as above, the presence of any of these constraint qualifications at $\bar{z} \in C$ is equivalent to the existence of *Lagrange multipliers*. That is, the existence of $\bar{\lambda} \in \mathbb{R}^p_+$ such that $\bar{\lambda}_j A_j(\bar{z}) = 0$ for $j = 1, \ldots, p$. We will say that such a $\bar{\lambda}$ is the Lagrange multiplier associated with the constraint $A(z) \leq 0$ at $\bar{z}$, or when clear with C, and define the sets

$$\begin{aligned} I_+(\bar{z}, \bar{\lambda}) &:= \left\{ j \in I(\bar{z}) \,\middle|\, \bar{\lambda}_j > 0 \right\} \text{ “strongly active set”} \\ I_0(\bar{z}, \bar{\lambda}) &:= \left\{ j \in I(\bar{z}) \,\middle|\, \bar{\lambda}_j = 0 \right\} \text{ “weakly active set”} \end{aligned}$$

We will often use the index sets defined here as subscripts to matrices as mentioned at the beginning of this chapter, leaving off the arguments when it is clear, e.g., $\bar{\lambda}$ is unique.

Finally, given $(\bar{x}, \bar{z}) \in \operatorname{gph} S$, where S is as described in the previous section, we say that the *strong second-order sufficient condition* (SSOSC) holds at $(\bar{x}, \bar{z})$, if

$$\langle h, \nabla_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda})h \rangle > 0 \qquad \begin{array}{l} \forall h \in \ker(\nabla A_{I_+(\bar{z},\bar{\lambda})}(\bar{z})) \setminus \{0\} \\ \forall \bar{\lambda} \in N_{\mathbb{R}^p_-}(A(\bar{z})) : \nabla^T A(\bar{z})\bar{\lambda} = -F(\bar{x}, \bar{z}), \end{array}$$

where $\mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) = F(\bar{x}, \bar{z}) + \nabla^T A(\bar{z})\bar{\lambda}$ represents the so-called *D-Lagrangian* (Outrata et al. [1998]) or *vector Lagrangian* (Luo et al. [1997]) and the vectors $\bar{\lambda} \in \mathbb{R}^p_+$ represent the Lagrange multipliers $\bar{\lambda}$ associated with C (as defined in preceeding section) at $\bar{z}$. Note that if there exists a differentiable function $g(x, z)$ with $\nabla_z g(x, z) = F(x, z)$. then setting $\mathcal{L}(x, z, \lambda) = 0$ along with the complementarity relations $0 \geq A(z) \perp \lambda \geq 0$ amounts to the classical Karush-Kuhn-Tucker (KKT) conditions associated with the following parametric non-linear program

$$\min_z \{g(x, z) \,|A(z) \leq 0\}.$$

Finally, we note that there are of course weaker second-order conditions (see e.g., Bonnans and Shapiro [2000]), however, this definition will suffice for our analysis. The version provided here has been taken from Robinson [1980], where it was noted as being a generalization of an older condition due to Fiacco and McCormick (see Chapter 2 in Fiacco and McCormick [1968]) and will be of particular interest in this paper, as it is used in combination with other constraints qualifications to guarantee properties such as strong regularity.

Chapter 3

Equilibrium Problems with Equilibrium Constraints

In this chapter, we formally define equilibrium problems with equilibrium constraints (EPECs) in both deterministic and stochastic settings. Well-posedness issues of EPECs are discussed and an EPEC modeling oligopolistic competition in an electricity spot market is introduced.

3.1 Mathematical Programs with Equilibrium Constraints

As noted in the introduction to this thesis, EPECs are formed via a coupling together of certain mathematical programs called mathematical programs with equilibrium constraints (MPECs). Using the standard references for MPECs (i.e., Luo et al. [1997] and Outrata et al. [1998]), we define an MPEC to be any optimization problem of the following form:

$$\min_{x,y} \{ f(x,y) \,|\, x \in X, y \in S(x) \} .$$

Here, $f : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}$, $S : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$, and $X \subseteq \mathbb{R}^s$ is non-empty. Typically, the multifunction S represents the solution mapping to a variational inequality or a generalized equation and we will usually require that f is continuously differentiable and X is closed or the entire space.

MPECs arise in a variety of real-life settings ranging from tax policy Hakonsen [April 1998] and traffic equilibrium Facchinei et al. [1996] to continuum mechanics Beremlijski et al. [2002, 2009] and represent an ideal mathematical formulation of Stackelberg, i.e., single-leader-multi-follower-type games. When $S(x)$ is merely a subset of $\mathbb{R}^t$ independent of x, then the MPEC reduces to a classical nonlinear optimization problem. Moreover, if $S(x)$ is the set of minimizers to some parametric optimization problem with parameter x and decision variable y, then the MPEC is called a *bi-level optimization problem* (see Dempe [2002]).

Though the notion of a solution for an MPEC coincides with the classical notion of a minimizer, the existence of even a local minimizer is not always clear. Such problems are due to complications arising from the equilibrium constraint. Nevertheless, there are some situations where the existence of a solution can be guaranteed for example Outrata et al. [1998, Propositions 1.1 and 1.2] using classical results and for more general settings we refer the reader to Loridan and Morgan [1989] and Zhang [1994].

Another difficulty involved in the analysis of MPECs appears in the derivation optimality conditions. The reason for this difficulty also stems from the equilibrium constraint, which causes the complete failure or meaninglessness of classical constraint qualifications. We are thus forced to use either weaker notions of stationarity than classical KKT-conditions or require that the MPEC of interest is endowed with a significant level of regularity. A more detailed discussion of stationarity conditions is provided in Chapter 4.

3.2 Equilibrium Problems with Equilibrium Constraints

Using the terminology of a two-level multi-leader-multi-follower game as discussed in the introduction, we can now formally define EPECs. On the upper level, we assume there exist n leaders, each of whom wishes to minimize its objective f_i, for all $i = 1, \ldots, n$, by using a strategy x_i from its set of *admissible strategies* $X_i \subseteq \mathbb{R}^s$. However the value of f_i is dependent not only on the decisions of all other leaders, represented here by the vector $x_{-i} := (x_1, \ldots x_{i-1}, x_{i+1}, \ldots, x_n)$, but also on the responses z of those players with a temporal disadvantage, i.e., the followers, whose decisions are modeled by an equilibrium constraint. Though a more general form of equilibrium constraint is possible, we will usually consider a setting where the lower-level decisions z are determined by the (perturbed) generalized equation

$$0 \in F(x, z) + N_C(z).$$

Here, $F : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}^t$ should be at least continuous and $C \subseteq \mathbb{R}^t$ is a closed convex set[1]. Unless otherwise noted, we denote the *solution mapping* of the generalized equation by $S(x)$, i.e., the multifunction that assigns to each $x \in X$ the solution(s) z to the generalized equation.

Because all players decisions are influenced by the reactions of the followers, the generalized equation couples together each of the optimization problems that the leaders are trying to solve as a common constraint. This n-tuple of mutually coupled optimization problems is an EPEC, which we denote by

$$\min_{x_i \in X_i, z} \{f_i(x_{-i}, x_i, z) \,|\, 0 \in F(x_{-i}, x_i, z) + N_C(z)\} \quad (i = 1, \ldots, n). \tag{3.1}$$

Here, $x = (x_{-i}, x_i)$ for all $i = 1, \ldots, n$ is used to remind the reader that part of x is a parameter and the other is a decision variable. Note that the feasibility condition $z \in C$ is implicitly implied via the normal cone mapping $N_C(z)$. Each of the coupled problems is then a standard MPEC in variables (x_i, z) parameterized by x_{-i} and we define for a fixed $\bar{x}_{-i}$ such that $(\bar{x}_{-i}, x_i)$ is an admissible strategy, $S_i(x_i) := S(\bar{x}_{-i}, x_i)$.

[1] The usage of a convex set C is merely to avoid determining which normal cone N_C is.

Formally, we say that a vector $(\bar{x}, \bar{z})$ is a (local) solution to (3.1), if for all $i = 1, \ldots, n$, the pair $(\bar{x}_i, \bar{z})$ belongs to the set of (local) solutions to the MPEC

$$\min_{x_i \in X_i, z} \{ f_i(\bar{x}_{-i}, x_i, z) \,|\, 0 \in F(\bar{x}_{-i}, x_i, z) + N_C(z) \}.$$

Such a solution is in line with the classical notion of a non-cooperative or Nash-Cournot equilibrium, only that here the leaders play a Stackelberg game with the followers as well. Thus it should also be clear that not every EPEC has a solution[2]. Yet the useful modeling framework they provide is undeniable and there are many real-life problems whose formal setting is best modeled by an EPEC, especially studies of market behavior see e.g., Hu [2002], Ehrenmann [2004a], Červinka [2008], and Su [2005].

3.3 Well-posedness Issues for EPECs

Using the terminology from Červinka [2008], the EPEC formulation (3.1) is referred to as *multioptimistic*. As noted, EPECs of this type may be ill-posed, i.e., if S_i is multivalued, it is not clear to the leaders which solution the followers will be taking. Throughout this section, we will consider EPECs of the following type

$$\min_{x_i \in X_i} \{ f_i(x_{-i}, x_i, z) \,|\, z \in S(x_{-i}, x_i) \} \;\; (i = 1, \ldots, n), \tag{3.2}$$

where $X_i \subseteq \mathbb{R}^s$, $f_i : \mathbb{R}^{ns+t} \to \mathbb{R}$, $S : \mathbb{R}^{ns} \rightrightarrows \mathbb{R}^t$. Given this setting, we define the following *best-response*-type mappings for each i^{th} MPEC

$$\Lambda_i(x) := \operatorname{argmin} \{ f_i(x, z) \,|\, z \in S(x) \} \;\; (i = 1, \ldots, n), \tag{3.3}$$

where the best response is that of the followers in reaction to any strategy of the leaders. Then clearly,

$$(\bar{x}, \bar{z}) \text{ is a solution to (3.2)} \Rightarrow \bar{z} \in \bigcap_{i=1}^{n} \Lambda_i(\bar{x}),$$

from which we can infer the following *necessary* condition for a solution to (3.2)

$$\text{(3.2) has a solution } \Rightarrow \exists x \in X : \bigcap_{i=1}^{n} \Lambda_i(x) \neq \emptyset. \tag{3.4}$$

The reverse direction, however, does not hold in general. To see this, consider the following example:

Example 3.1 (failure of sufficiency of the optimality condition). *For the EPEC (3.2), let $n = s = t = 2$, $X_i := [0, 1]$ for $i = 1, 2$, $S(x) := \mathbb{B}_r(x)$, $r > 0$, where $r \neq 1/2$; define $f_1(x, z) := 1 - x_2 - z_2$ and $f_2(x, z) := 1 - x_1 - z_1$. One can*

[2] Even very simple non-cooperative games may have no solution, see e.g., Osborne and Rubinstein [1994, Exercise 17.1]

then deduce that

$$\Lambda_1(1,1) \cap \Lambda_2(1,1) = \{(1+r, 1+r)\}.$$

That is, $\exists (x_1, x_2) \in X$ *such that* $\Lambda_1(x_1,x_2) \cap \Lambda_2(x_1,x_2) \neq \emptyset$. *Nevertheless, consider for any* $(\bar{x}_1, \bar{x}_2) \in [0,1]^2$ *that*

$$\underset{\substack{x_1 \in [0,1] \\ z}}{\operatorname{argmin}} \{1 - \bar{x}_2 - z_2 \,|\, z \in S(x_1, \bar{x}_2)\} = [0,1] \times [[-r, 1+r] \times \{\bar{x}_2 + r\}]$$

$$\underset{\substack{x_2 \in [0,1] \\ z}}{\operatorname{argmin}} \{1 - \bar{x}_1 - z_1 \,|\, z \in S(\bar{x}_1, x_2)\} = [0,1] \times [\{\bar{x}_1 + r\} \times [-r, 1+r]]$$

Though it seems that $(\bar{x}_1, \bar{x}_2, \bar{x}_1 + r, \bar{x}_2 + r)$ *would be a solution to the EPEC, because*

$$S(x) = \mathbb{B}_r(x_1, x_2) := \left\{ z \,\middle|\, (z_1 - x_1)^2 + (z_2 - x_2)^2 \leq r^2 \right\},$$

substituting $(\bar{x}_1, \bar{x}_2, \bar{x}_1 + r, \bar{x}_2 + r)$ *into this inequality yields* $2r^2 \leq r^2$. *Therefore, this EPEC has* ***no*** *solutions.*

Example 3.1 illustrates some of the possible problems one may encounter when modeling a situation via an EPEC. However, there are some conditions we can impose ensuring that the EPEC will in fact be well-posed without requiring that S is always single-valued.

Proposition 3.1 (**sufficient conditions for well-posedness**). *If there exists a subset* $J \subseteq \{1, \ldots, t\}$ *and a neighborhood* $\mathcal{U} \subseteq X$ *such that the following two conditions hold*

1. $S(x) \neq \emptyset$ *and* $\Pi_J(S(x))$ *is a singleton for all* $x \in \mathcal{U} \subset X$

2. *The objective functions* f_i *are independent of* z_j *for* $j \in J^c$, *for all* $i = 1, \ldots, n$,

then

$$\bigcap_{i=1}^{n} \Lambda_i(x) = \Pi_J(S(x)) \times \Pi_{J^c}(S(x)) \neq \emptyset, \quad \forall x \in \mathcal{U}$$

Here, Π_J *represents the projection onto the subspace* $\operatorname{span}\{e^i | i \in J\}$, *where* e^i *is the standard unit vector in* $\mathbb{R}^t$

Proof. Fix arbitrary $J \subseteq \{1, \ldots, t\}$ and $\mathcal{U}$ both satisfying 1. and 2., and let $z := (z_J, z_{J^c})$. Consider for some $\hat{x} \in \mathcal{U} \subseteq X$,

$$\Lambda_i(\hat{x}) = \operatorname{argmin} \{f_i(\hat{x}, z) \,|\, z \in S(\hat{x})\}.$$

Then

$$z \in S(\hat{x}) \Leftrightarrow z_J \in \Pi_J(S(\hat{x})),\ z_{J^c} \in \Pi_{J^c}(\Pi_J^{-1}(z_J) \cap S(\hat{x})).$$

This implies via 2 that

$$\Lambda_i(\hat{x}) = \operatorname{argmin} \left\{ f_i(\hat{x}, z_J) \,\middle|\, (z_J, z_{J^c}) \in \Pi_J(S(\hat{x})) \times \Pi_{J^c}(\Pi_J^{-1}(z_J) \cap S(\hat{x})) \right\},$$

which due to assumption 1. leads to

$$\Lambda_i(\hat{x}) = \\ \operatorname{argmin} \left\{ f_i(\hat{x}, \Pi_J(S(\hat{x}))) \,\middle|\, (z_J, z_{J^c}) \in \Pi_J(S(\hat{x})) \times \Pi_{J^c}(\Pi_J^{-1}(z_J) \cap S(\hat{x})) \right\}.$$

But then $f_i(\hat{x}, \Pi_J(S(\hat{x})))$ is a constant and

$$\bigcap_{i=1}^{n} \Lambda_i(\hat{x}) = \Pi_J(S(\hat{x})) \times \Pi_{J^c}(S(\hat{x})) \neq \emptyset,$$

as $S(\hat{x}) \neq \emptyset$ for all $\hat{x} \in \mathcal{U}$. □

Notice how Example 3.1 violates the assumptions of Proposition 3.1, being that for $J = \{1\}, \{2\}$ and $\{1, 2\}$ and any feasible x, $\Pi_J(S(x_1, x_2))$ is not a singleton. Furthermore, $J = \emptyset$ would implies that f_1 and f_2 cannot depend on z, which is also not the case.

3.4 EPEC modeling Oligopolistic Competition in an Electricity Spot Market

We now present an EPEC that models oligopolistic competition in an electricity spot market. The model used here is inspired by work seen in Borenstein et al. [2000], Escobar and Jofre [2005], Hu and Ralph [2005], Hu et al. [2007] and has recently been investigated in Henrion and Römisch [2007].

Assume that the network of interest is represented by a *connected* oriented graph with m edges (transmission lines) and $N > 1$ nodes. Within the framework of the spot market model, $B \in \mathbb{R}^{N\times m}$ is used to represent the incidence matrix of the electricity network, with entries

$$b_{ij} = \begin{cases} 1 & \text{if edge } j \text{ enters node } i \\ -1 & \text{if edge } j \text{ leaves node } i \\ 0 & \text{otherwise} \end{cases}$$

The connectedness of the graph implies that for all $i = 1, \ldots, N$, there exists at least one index $j \in \{1, \ldots, m\}$ such that b_{ij} equals 1 or -1. Certainly if this were not true, then B would contain a row of zeros, thereby indicating that the node whose index corresponds to this row is isolated from the rest of the graph. Finally, we assume that at each node, electricity is both in demand and generated and that generator i exists only at node i, for all $i = 1, \ldots, N$.

As noted in Escobar and Jofre [2005], it is reasonable to use the following mapping to represent the amount of electricity lost due to transmission

$$L(y) = \left(\tfrac{1}{2} \textstyle\sum_{j=1}^{m} |b_{1j}| \rho_j y_j^2, \ldots, \tfrac{1}{2} \sum_{j=1}^{m} |b_{Nj}| \rho_j y_j^2 \right)^T. \tag{3.5}$$

Here, $y \in \mathbb{R}^m$ denotes the oriented flow vector of electricity along the edges of the graph and $\rho_j \geq 0$ is the loss coefficient of line j, for all $j = 1, \ldots, m$. Let

$i, k \in \{1, \dots, N\}$, with $i \neq k$, be two nodes connected by edge $j \in \{1, \dots, m\}$. If $\rho_j = 0$, then we interpret this to mean that generators i and k are reasonably geographically close and thus the loss of electricity due to transmission is considered negligible. We will observe in Chapter 8 that by setting all $\rho_j = 0$ $(j = 1, \dots, m)$ we can obtain valuable qualitative and quantitative information about certain solutions.

Given these considerations, we model the satisfaction of demand with the following system of inequalities

$$q + By \geq d + L(y). \tag{3.6}$$

Here, the parameter $d \in \mathbb{R}^N$ represents the vector of demands at each node and $q \in \mathbb{R}^N$ is the vector of electricity generated at each respective node. Finally, we have the following bounds on production and flow

$$0 \leq q_i \leq \hat{q}_i \quad (i = 1, \dots, N) \qquad -\hat{y}_j \leq y_j \leq \hat{y}_j \quad (j = 1, \dots, m).$$

Electricity spot market models are structured in such a way so that each of the competing firms, or generators as we refer to them here, bids a quadratic cost function to an independent system operator (ISO)

$$c_i(\alpha_i, \beta_i, q_i) = \alpha_i q_i + \beta_i q_i^2 \quad (i = 1, \dots, N).$$

Nevertheless, the bid linear and quadratic cost coefficients α_i and β_i may in reality differ from the true cost coefficients γ_i and δ_i, respectively. Yet it is assumed that neither the ISO nor the other generators know their competitors true cost coefficients, hence, the ISO determines generation and flow such that demand is met in each node of the network and that the overall costs are minimized given the bid cost functions $c_i(\alpha_i, \beta_i, q_i)$

$$\min_{q,y} \left\{ \sum_{i=1}^{N} c_i(\alpha_i, \beta_i, q_i) \mid (q, y) \in G \right\}, \tag{3.7}$$

where

$$G := \left\{ (q, y) \in \mathbb{R}^{N+m} \mid q + By \geq d + L(y),\ 0 \leq q \leq \hat{q},\ -\hat{y} \leq y \leq \hat{y} \right\}.$$

It should be noted that the vector (α, β) appears only as a perturbation parameter in (3.7) and is therefore not considered a decision variable on this level. The parametric optimization problem (3.7) is referred to as the ISO or dispatch problem. Moreover, for certain $(\alpha', \beta') \in \mathbb{R}^{2N}$, it is clear that the objective function is convex, in some cases even strongly convex. In such cases, we know that the corresponding optimal solutions (q', y') are characterized as solutions of the following generalized equation arising from the KKT conditions of (3.7)

$$0 \in \begin{pmatrix} \alpha' + 2[\operatorname{diag} \beta']q \\ 0 \end{pmatrix} + N_G(q, y). \tag{3.8}$$

Here, we use $[\operatorname{diag}\beta']$ to denote the diagonal matrix with entries β'_i along the diagonal.

In order to derive each generator's profit function, we need what is known as the clearing price function. Given any generator's bid cost function and taking the derivative with respect to q_i, one obtains this function, the values of which indicate the price at which quantity supplied equals quantity demanded

$$\pi_i(q_i) := \nabla_{q_i} c_i(\alpha_i, \beta_i, q_i) = \alpha_i + 2\beta_i q_i.$$

Using the coefficients from $\pi_i(q_i)$ along with the corresponding true linear and quadratic coefficients, we define the i^{th} profit function

$$f_i(\alpha_i, \beta_i, q, y) := (\alpha_i - \gamma_i)q_i + (2\beta_i - \delta_i)q_i^2.$$

Since the solution type sought is a non-cooperative equilibrium, each generator i must solve the following MPEC, which is formed from (3.8) by fixing the decisions of all other competitors

$$\max_{\substack{(\alpha_i,\beta_i)\in\mathbb{R}^2\\ q,y}} \left\{ f_i(\alpha_i, \beta_i, q, y) \middle| 0 \in \begin{pmatrix} \zeta(\alpha_i, \beta_i, q) \\ 0 \end{pmatrix} + N_G(q, y) \right\} \quad (i = 1, \ldots, N),$$

where $\zeta(\alpha_i, \beta_i, q) := (\bar{\alpha}_{-i}, \alpha_i) + 2[\operatorname{diag}(\bar{\beta}_{-i}, \beta_i)]q$. Then the coupled system of MPECs sharing the same equilibrium constraint represents an EPEC:

$$\min_{\substack{(\alpha_i,\beta_i)\in\mathbb{R}^2\\ q,y}} \left\{ -f_i(\alpha_i, \beta_i, q, y) \middle| 0 \in \begin{pmatrix} \alpha + 2[\operatorname{diag}\beta]q \\ 0 \end{pmatrix} + N_G(q, y) \right\},$$

with $i = 1, \ldots, N$. We switch here from maximization of profit to minimization of negative profit merely for the fact that the stationarity conditions which we will use are formulated for EPECs defined via minimization problems. For notational simplicity, we define

$$F(\alpha, \beta, q, y) := \begin{pmatrix} \alpha + 2[\operatorname{diag}\beta]q \\ 0 \end{pmatrix}.$$

and we can thus rewrite our spot market EPEC in the compact form

$$\min_{\substack{(\alpha_i,\beta_i)\in\mathbb{R}^2\\ q,y}} \{ -f_i(\alpha, \beta, q, y) \,|0 \in F(\alpha, \beta, q, y) + N_G(q, y) \} \; (i = 1, \ldots, N). \qquad (3.9)$$

Proposition 3.2 (well-posedness of the spot market EPEC). *If $G \neq \emptyset$, then the spot market EPEC is well-posed.*

Proof. We will use the criteria outlined by Proposition 3.1 to show the well-posedness. Choose $\mathcal{U} \subseteq \mathbb{R}^{2N}$ such that the object function to (3.7) is strongly convex[3] in q and denote by $S(\alpha, \beta)$ the solution mapping to (3.8). Then clearly

$$S(\alpha, \beta) = \operatorname*{argmin}_{q,y} \left\{ \sum_{i=1}^{N} c_i(\alpha_i, \beta_i, q) \,|\, (q, y) \in G \right\} \neq \emptyset, \ \forall (\alpha, \beta) \in \mathcal{U},$$

where the last equality follows from G being non-empty and convex. As the generator's objective functions do not depend on y, we need to show that the projection onto the q component, denoted $\Pi_q(S(\alpha, \beta))$ is a singleton for all $(\alpha, \beta) \in \mathcal{U}$. Fix an arbitrary $(\bar{\alpha}, \bar{\beta}) \in \mathcal{U}$ and define $\widetilde{G} := \{q \,|\, \exists y : (q, y) \in G\}$ along with the minimization problem:

$$\min_{q} \left\{ \sum_{i=1}^{N} c_i(\bar{\alpha}_i, \bar{\beta}_i, q) \,\middle|\, q \in \widetilde{G} \right\}. \tag{3.10}$$

Clearly, $\widetilde{G} \neq \emptyset$ and convex. Then because the objective function of (3.10) is strongly convex, it has a unique minimizer $\tilde{q} \in \widetilde{G}$, i.e., $\forall q \in \widetilde{G}$

$$\sum_{i=1}^{N} c_i(\bar{\alpha}_i, \bar{\beta}_i, \tilde{q}) \leq \sum_{i=1}^{N} c_i(\bar{\alpha}_i, \bar{\beta}_i, q). \tag{3.11}$$

Since $q \in \widetilde{G} \Leftrightarrow \exists y : (q, y) \in G$, there exists a y such $(\tilde{q}, y) \in G$ and (3.11) holds $\forall (q, y) \in G$ as well. Assume now that there exists $(\hat{q}, \hat{y}) \in G$ such that $(\hat{q}, \hat{y})$ is a minimizer of (3.7) and $\hat{q} \neq \tilde{q}$. But by definition, $\hat{q} \in \widetilde{G}$ and (3.11) holds for all $q \in \widetilde{G}$ as well, contradicting the uniqueness of $\tilde{q}$ due to the strong convexity of the objective function of (3.10). Therefore $\Pi_q(S(\alpha, \beta))$ is a singleton for all $(\alpha, \beta) \in \mathcal{U}$ □

The requirement in Proposition 3.2 that the feasible set be non-empty amounts essentially to the existence of a feasible flow. This can be argued via the well-known Gale-Hoffman inequalities (cf. Rockafellar [1984]) by making some assumptions on the production and flow capacities.

3.5 Stochastic MPECs

In order to construct stochastic EPECs, we need to define their basic components, namely, stochastic MPECs (SMPECs). Beginning in the mid-twentieth century, the study of optimization problems with random parameters, i.e., *stochastic programs*, has become a vibrant area of research involving elements of mathematical programming, statistics, variational and functional analysis. Stochastic programs enjoy a variety of applications, particularly in finance, economics, and logistics and their study presents the researcher with both theoretical as well as numerical difficulties not always present in the study of finite-dimensional optimization problems, e.g., integrability of feasible solutions and dimensionality. A few references on the basics and advances in stochastic programming can be found in

[3] E.g., define $\mathcal{U}$ such that $\alpha_i, \beta_i > 0$ for all $i = 1, \ldots, N$ and $\alpha, \beta \in \mathcal{U}$

Birge and Louveaux [1997] and Ruszczyński and Shapiro [2003], respectively.

The study of SMPECs is a relatively new addition to stochastic programming with the formal definition and statements concerning the existence of optimal solutions to SMPECs first appearing in Patriksson and Wynter [1999] and Evgrafov and Patriksson [2004]. Some studies in the context of the implicit programming approach (cf. Luo et al. [1997, Chapter 4] or Outrata et al. [1998, Chapter 7]) can be found in the papers Xu [2005] and Xu [2006]. More recently, there have been studies carried out concerning the basic properties of SMPECs in the framework of so-called 'here-and-now' type stochastic programs Shapiro [2006], Shapiro and Xu [2008] and the development of optimality conditions in situations where the lower-level solutions are non-unique can be found in Xu and Ye [2009].

Though there are perhaps a few ways to write an SMPEC, we choose to use the form

$$\min_{x,z(\cdot)} \left\{ \int_\Omega f(x,z,\omega) d\mathbb{P}(\omega) \middle| x \in X, z \in S(x,\omega) \right\}. \tag{3.12}$$

Here, we let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space and assume that $f : \mathbb{R}^s \times \mathbb{R}^t \times \Omega \to \mathbb{R}$, and $S : R^s \times \Omega \rightrightarrows \mathbb{R}^t$. Note that we write $z(\cdot)$ to indicate the dependence on ω.

In principal, one can consider an SMPEC to be a certain type of parametric optimization problem in the sense of those studied in Bank et al. [1982] and Bonnans and Shapiro [2000], only that here, the parameter is a probability measure, thus making their study quite difficult due to the complicated nature and oft lack of regularity in the accompanying parameter space.

3.6 Stochastic EPECs

Though there have been extensions of the original models of Sherali, Soyster, and Murphy in the stochastic setting, see e.g., Wolf and Smeers [1997] and DeMiguel and Xu [2009], these models do qualify as EPEC per se, as they do not contain equilibrium constraints. In truth, under the moniker *stochastic EPEC* or SEPEC only one published work is available, namely Henrion and Römisch [2007], in which true SEPECs are considered.

As in the deterministic setting, we analogously define an SEPEC by coupling together n SMPECs. Using the same data assumptions as in the definition of (3.12), we denote a SEPEC as follows

$$\min_{x_i,z(\cdot)} \left\{ \int_\Omega f_i(x_{-i}, x_i, z, \omega) d\mathbb{P}(\omega) \middle| x_i \in X_i, z \in S(x_{-i}, x_i, \omega) \right\} \quad (i = 1, \dots, n). \tag{3.13}$$

As we assume the random parameter enters the models on the lower level, we do not need different random variables for each of the SMPECs making up the SEPEC. It is also important to note that the solution of a SEPEC is the same as that for an EPEC and should not be confused with a classical mixed-strategy equilibrium. The difference from EPEC to SEPEC solutions lies merely in the fact that part of the solution, i.e., the part depending implicitly on the random parameter(s), exists in function space.

3.7 A Stochastic Spot Market EPEC

We now present the spot market model with stochastic demand d. In reality, no generator i can truly know the demand d_i associated with its respective node-or at the very least, it does not know the demands d_j, for $j \neq i$. However, in the event that historical data is available to the generators, it is common practice to assume that d is some random vector from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into $\mathbb{R}^N$ whose distribution is at least approximately known. Note that we will use ω as an argument to indicate dependence upon the random parameter. Such an assumption allows one to define the following closed- and convex-valued multifunction $G : \Omega \rightrightarrows \mathbb{R}^{N+m}$:

$$G(\omega) := \left\{ (q, y) \in \mathbb{R}^{N+m} \middle| \, q + By \geq d(\omega) + L(y),\ 0 \leq q \leq \hat{q},\ -\hat{y} \leq y \leq \hat{y} \right\}.$$

Thus, the pair (q, y) of generation and flow becomes (implicitly) an $\mathbb{R}^{N+m}$-valued random vector-valued function on $(\Omega, \mathcal{F}, \mathbb{P})$, whereby the ISO minimizes the overall costs, i.e.,

$$\min_{q,y} \left\{ \sum_{i=1}^{N} c_i(q_i(\omega)) \,|(q(\omega), y(\omega)) \in G(\omega), \mathbb{P}\text{-a.s.} \right\} \tag{3.14}$$

Under the assumption that the mappings $(q(\cdot), y(\cdot))$ lie in a Banach space [4] (cf. Bonnans and Shapiro [2000, Chapter 3, Section 1]), the first-order optimality conditions of (3.14) become

$$0 \in \begin{pmatrix} \alpha + 2\,[\operatorname{diag}\beta]\, q(\omega) \\ 0 \end{pmatrix} + N_{G(\omega)}(q(\omega), y(\omega)),\ \mathbb{P}\text{-a.s.},$$

where we assume $G(\omega) \neq \emptyset$ almost everywhere. The generators then wish to maximize their expected profit, or equivalently minimize their expected losses. Therefore, the EPEC (3.9) is transformed into the following stochastic equilibrium problem with equilibrium constraints (SEPEC)

$$\begin{aligned} &\min_{\substack{\alpha_i,\beta_i \\ q(\cdot),y(\cdot)}} \left\{ \int_\Omega \Big((\gamma_i - \alpha_i)\, q_i(\omega) + (\delta_i - 2\beta_i)\, q_i^2(\omega) \Big)\, d\mathbb{P}(\omega) \,\middle|\, \right. \\ &\left. 0 \in \begin{pmatrix} \alpha + 2\,[\operatorname{diag}\beta]\, q(\omega) \\ 0 \end{pmatrix} + N_{G(\omega)}(q(\omega), y(\omega)),\ \mathbb{P}\text{-a.s.} \right\} \ (i = 1, \ldots, N), \end{aligned} \tag{3.15}$$

where the pairs (α_i, β_i), $i = 1, \ldots, N$, are deterministic and have to be determined before the realization of the demand, and the pairs $(q_i(\cdot), y_i(\cdot))$ $i = 1, \ldots, N$, are stochastic. In the terminology of two-stage stochastic programming with recourse,

[4]We will see in the proof of Proposition 3.3, that $(q, y) \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}) \times \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$,i.e., the space of real-valued 2-dimensional random vectors with finite second moments.

the cost coefficients (α_i, β_i) are first-stage decisions, while $(q_i(\cdot), y_i(\cdot))$ are second-stage or recourse decisions.

Next, we show that the SEPEC (3.15) is well-defined, i.e., the objective function is integrable at a solution[5]. In order to do so, we will need the following definitions.

Definition 3.1. *A single-valued mapping $h : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ is called a* ***Carathéodory mapping*** *when $h(x, \omega)$ is measurable in ω for each fixed x and continuous in x for each fixed ω. A function $g : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ is called a* ***normal integrand*** *if its epigraphical mapping* $\operatorname{epi} g(\cdot, \omega) := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \,|\, g(x, \omega) \leq \alpha\}$ *is closed-valued and measurable. x is a* ***measurable selection*** *of a multivalued function $S : \Omega \rightrightarrows \mathbb{R}^n$ if $x : \operatorname{dom} S \to \mathbb{R}^n$ such that $x(\omega) \in S(\omega)$ for all $\omega \in \Omega$.*

Proposition 3.3 (**well-definedness of the expected value functional**)**.** *If $G(\omega) \neq \emptyset$ $\mathbb{P}$-a.s, then stochastic spot market SEPEC (3.15) is well-defined.*

Proof. Clearly each function of the form $g_i(q, y, \omega) := d_i(\omega) + L_i(y) - q - B_i y$ is a normal integrand, thus by Rockafellar and Wets [1998, Theorem 14.36], $G(\omega)$ is closed-valued and measurable and therefore admits a measurable selection $(q(\cdot), y(\cdot)) : \Omega \to \mathbb{R}^{N+m}$. Now for fixed (α, β), continuity of the objective function in (3.7) implies measurability in ω for any measurable selection. Now fix some $\omega' \in \Omega$. By assumption, $G(\omega') \neq \emptyset$ $\mathbb{P}$-a.s. Then since the objective function of (3.14) is continuous over $G(\omega')$, it is Carathéodory. Thus, as per Rockafellar and Wets [1998, Theorem 14.37], the multifunctions $\Psi : \Omega \rightrightarrows \mathbb{R}^{N+m}$ defined:

$$\Psi(\omega) := \operatorname*{argmin}_{q,y} \left\{ \sum_{i=1}^{N} c_i(\alpha, \beta, q_i(\omega)) \,|\, (q(\omega), y(\omega)) \in G(\omega), \mathbb{P}\text{-a.s.} \right\}$$

is closed valued and measurable. Then by Rockafellar and Wets [1998, Theorem 14.6], $\Psi(\omega)$ admits a measurable selection: $(q(\omega), y(\omega)) \in \Psi(\omega)$. Since $0 \leq q_i(\omega) \leq \hat{q}$ $\mathbb{P}$-a.s., for all $i = 1, \ldots, N$, we have:

$$(\gamma_i - \alpha_i) \int_\Omega q_i(\omega) d\mathbb{P}(\omega) + (\delta_i - 2\beta_i) \int_\Omega q_i(\omega)^2 d\mathbb{P}(\omega) < \infty$$

as was to be shown. □

Remark 3.1. *Note that though we have extensively used Theorems from Chapter 14 of Rockafellar and Wets [1998], though it should be noted that many of these results can be found in the much older book book Castaing and Valadier [1977] as well as in Aubin and Frankowska [1990, Chapters 8 and 9]*

[5] A well-posedness argument similar to that from Proposition 3.2 could also be applied

Chapter 4

Dual Stationarity Concepts for MPECs and EPECs

In this chapter, we present certain dual concepts of stationarity important to the study of MPECs and EPECs. Though there are also primal notions of stationarity, e.g., *Bouligand* or *B-stationarity* (introduced in terms of MPECs in Luo et al. [1996]), which utilize contingent cones and directional derivatives, we are primarily interested in using the dual concepts. Conditions of this type are beneficial in two ways. First, many numerical procedures are developed using dual stationarity conditions, e.g., Leyffer and Munson [April 2005]. Second, the multipliers arising from such conditions allow us to better characterize solutions. Since we will later observe that stationarity conditions for EPECs are composed of the individual stationarity conditions for the MPECs making up the EPEC in question, the main results and definitions are presented in terms of MPECs, after which we discuss how these results can be extended to EPECs. In this sense, this chapter provides new explicit multiplier-based stationarity conditions for both MPECs and EPECs.

4.1 Strong Stationarity

Throughout the following sections, we will consider MPECs of the type:

$$\min_{x,z} \{ f(x,z) \,|\, x \in X, z \in S(x) \} . \tag{4.1}$$

Here, $f : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}$ is continuously differentiable, $X \subseteq \mathbb{R}^s$ is non-empty and closed, and $S : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$ is the solution mapping to the generalized equation

$$0 \in F(x,z) + N_C(z), \tag{4.2}$$

where $F : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}^t$ is continuously differentiable and

$$C := \left\{ z \in \mathbb{R}^t \,|\, A_j(z) \leq 0, j = 1, \ldots, p \right\},$$

with $A_j(z)$ twice continuously differentiable and convex for all $j = 1, \ldots, p$. We let $A(z) := (A_1(z), \ldots, A_p(z))^T$ and note that the convexity assumption implies $N_C(z)$ is the standard normal cone from convex analysis.

Given this setting, we begin by providing the definition of the strongest (dual) stationarity concept for MPEC solutions.

Definition 4.1 (**strong stationarity**). *A feasible point* $(\hat{x}, \hat{z})$ *to (4.1) is called* ***strongly*** *or* ***S-stationary*** *if the following condition holds*

$$0 \in \nabla f(\hat{x}, \hat{z}) + \widehat{N}_{\operatorname{gph} S \cap [X \times \mathbb{R}^t]}(\hat{x}, \hat{z}). \tag{4.3}$$

We refer to (4.3) as the S-stationarity conditions associated with the MPEC (4.1). In the framework of MPECs, some references important to the analysis and derivation of such conditions include the monograph Luo et al. [1997] and Pang and Fukushima [1999], in which they were defined via the dualization of B-stationarity conditions. Their connections to other weaker stationarity concepts as well as conditions enabling a more explicit version of (4.3) have been detailed in Ye [1999], Scheel and Scholtes [2000], Flegel et al. [2007].

S-stationarity conditions amount to the classical first-order dual optimality conditions for the problem of minimizing a smooth function over a closed domain when $(\hat{x}, \hat{z})$ is a local solution to (4.1) (see e.g., Rockafellar and Wets [1998, Theorem 6.12]) and provide the most selective form of stationarity conditions. Furthermore, if (4.1) can be expressed as a classical non-linear program, then (4.3) amounts to the corresponding KKT-conditions.

The real difficulty in obtaining usable conditions for characterizing solutions from (4.3) lies in the calculation of $\widehat{N}_{\operatorname{gph} S \cap [X \times \mathbb{R}^t]}(\bar{x}, \bar{z})$. This is a particularly non-trivial task due to the insufficient calculus of the Fréchet variational objects. Nevertheless, we are able to partially address this issue in the later sections of this chapter, thus allowing us to make a comparison of the selectivity and information provided by our two main stationarity concepts using some small examples of the spot market EPEC in Chapter 7 (see Examples 8.1 and 8.2).

4.2 M-Stationarity

As mentioned in the previous section, the usefulness of S-stationarity conditions in the form provided in (4.3) may be quite limited due to the difficulties in calculating the Fréchet normal cone. Therefore, we opt to use the next strongest form of stationarity conditions, which in particular make use of the limiting variational objects due to B.S. Mordukhovich.

Definition 4.2 (**M-stationarity**). *A feasible point* $(\hat{x}, \hat{z})$ *to (4.1) is called* ***Mordukhovich*** *or* ***M-stationary*** *if the following condition holds*

$$0 \in \nabla f(\hat{x}, \hat{z}) + N_{\operatorname{gph} S \cap [X \times \mathbb{R}^t]}(\hat{x}, \hat{z}).$$

M-stationarity conditions in this form first arose in Mordukhovich [1976] and Mordukhovich [1980]. In Ye and Ye [1997], calmness of a certain multifunction was used as a constraint qualification enabling the simplification of the limiting normal cone. This was elaborated on further in Ye [1999] for a more general class of MPECs. Two more essential references for M-stationarity conditions include Outrata [1999, 2000], in which careful attention was paid to verifying the calmness condition via the Aubin property and strong regularity as well as the explicit calculation of the coderivatives. For results on M-stationarity conditions in MPEC settings more similar to ours, we direct the reader to Flegel et al. [2007].

Recall from Chapter 2 that $\widehat{N}_\Gamma \subseteq N_\Gamma$ for any closed set Γ. Then it is easy to see from the previous definition, that any S-stationary point is an M-stationary point, but not necessarily every M-stationary point is an S-stationary point. Therefore, one must be careful when working with M-stationary points, as there is a higher probability that a given M-stationary point is not a solution. We will observe this phenomenon later in Chapter 7, Example 8.1.

As a testament to the complete calculus enjoyed by the limiting variational objects, we have the following theorem for a locally optimal solution to (4.1), (see Ye and Ye [1997, Theorem 3.2]).

Theorem 4.1 (**M-stationarity conditions using calmness**). *Let $(\bar{x}, \bar{z})$ be a local solution to (4.1). If the multifunction*

$$\Psi(u) := \Big\{(x, z) \in \mathbb{R}^s \times \mathbb{R}^t \,|u \in F(x, z) + N_C(z)\Big\} \tag{4.4}$$

is calm at $(0, \bar{x}, \bar{z})$, then there exists multipliers $v^* \in \mathbb{R}^t$ such that

$$0 \in \nabla_x f(\bar{x}, \bar{z}) + \nabla_x^T F(\bar{x}, \bar{z})v^* + N_X(\bar{x}) \tag{4.5}$$

$$0 \in \nabla_z f(\bar{x}, \bar{z}) + \nabla_z^T F(\bar{x}, \bar{z})v^* + D^*N_C(\bar{z}, -F(\bar{x}, \bar{z}))(v^*). \tag{4.6}$$

Note that (4.5) and (4.6) differ from the original M-stationarity condition in that the normal cone $N_{\operatorname{gph} S\cap[X\times\mathbb{R}^t]}(\hat{x}, \hat{z})$ has been replaced by an upper approximation (see Ye and Ye [1997, Proof of Theorem 3.2]). Nevertheless, we we will refer to solutions or feasible points $(\bar{x}, \bar{z})$ for which there exist $v^* \in \mathbb{R}^t$ such that (4.5) and (4.6) hold as M-stationary points.

Recent developments in the study of explicit formulae for coderivatives of the type displayed in (4.6) allow us to ultimately rewrite (4.5) and (4.6) even more explicitely, i.e., without any ambiguous terms, provided C enjoys certain regularity properties. Thus, the previous result often yields a significantly more workable set of stationarity conditions than those obtained via strong stationarity, albeit possibly weaker. We will discuss *transformation formulae* for coderivatives of normal cone mappings in Chapter 6, however, we make use of some of these formulae in the following sections.

Finally, we also note that Theorem 4.1 stands as the motivation for Chapter 5, i.e., the need for analytical conditions enabling the verification of stability properties of multifunctions of type (4.4); we call these multifunctions the *perturbation mappings* associated with the MPEC (4.1).

Corollary 4.1 (**explicit M-stationarity conditions I**). *Let $(\bar{x}, \bar{z})$ be a local solution to (4.1) and without loss of generality[1] assume that $A(\bar{z}) = 0$. Assume*

[1] It can be shown that inactive constraints play essentially no role in the calculation as the normal cones are considered locally. Moreover, in the application, we show that all such constraints are active at the considered solutions

1. *LICQ holds at* $\bar{z} \in C$

2. *The perturbation mapping (4.4) is calm at* $(0, \bar{x}, \bar{z})$

Then there exists a unique $\bar{\lambda} \in \mathbb{R}^p_+$ *and vectors* $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ *such that*

$$-\nabla_x f(\bar{x}, \bar{z}) \in \nabla_x^T F(\bar{x}, \bar{z}) v^* + N_X(\bar{x}) \tag{4.7}$$

$$-\nabla_z f(\bar{x}, \bar{z}) = \nabla_z^T F(\bar{x}, \bar{z}) v^* + \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) w^* \tag{4.8}$$

$$\nabla A_j(\bar{z}) v^* = 0 \quad \forall j : \bar{\lambda}_j > 0 \tag{4.9}$$

$$w_j^* = 0 \quad \forall j : \bar{\lambda}_j = 0,\ \nabla A_j(\bar{z}) v^* < 0 \tag{4.10}$$

$$w_j^* \geq 0 \quad \forall j : \bar{\lambda}_j = 0,\ \nabla A_j(\bar{z}) v^* > 0 \tag{4.11}$$

$$F(\bar{x}, \bar{z}) = -\nabla^T A(\bar{z}) \bar{\lambda} \tag{4.12}$$

Proof. Given 2., there exists $v^* \in \mathbb{R}^t$ such that (4.5) and (4.6) hold, from which we immediately obtain (4.7). Turning now to the coderivative in (4.6), 1. allows us to invoke Theorem 6.2, which states

$$D^* N_C(\bar{z}, -F(\bar{x}, \bar{z}))(v^*) = \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z}) v^*),$$

where $\bar{\lambda} \in \mathbb{R}^p_+$ is uniquely defined by (4.12). Referring now to Proposition 6.1, and substituting in $\mathcal{C} := \mathbb{R}^p_-$, $\bar{z} := A(\bar{z})$, $\bar{v} := \bar{\lambda}$, and $v^* := \nabla A(\bar{z}) v^*$ in order calculate the remaining coderivative; we see that the coderivative $D^* N_C(\bar{z}, -F(\bar{x}, \bar{z}))(v^*)$ can be replaced by the second order term in the previous relation plus $\nabla^T A(\bar{z}) w^*$ with v^* and $w^* \in \mathbb{R}^p$ satisfying (4.9)-(4.11). $\square$

We can also use another result from Chapter 6 to obtain a similar set of explicit M-stationarity conditions, albeit larger,i.e., weaker.

Corollary 4.2 (explicit M-stationarity conditions II). *Let* $(\bar{x}, \bar{z})$ *be a local solution to (4.1) and without loss of generality assume that* $A(\bar{z}) = 0$. *If*

1. *MFCQ holds at* $\bar{z} \in C$

2. *CRCQ holds at* $\bar{z} \in C$

3. *The multifunction defined in (4.4) is calm at* $(0, \bar{x}, \bar{z})$

Then there exist vectors $\bar{\lambda} \in \mathbb{R}^p_+$ *and* $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ *such that*

$$-\nabla_x f(\bar{x}, \bar{z}) \in \nabla_x^T F(\bar{x}, \bar{z}) v^* + N_X(\bar{x}) \tag{4.13}$$

$$-\nabla_z f(\bar{x}, \bar{z}) = \nabla_z^T F(\bar{x}, \bar{z}) v^* + \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) w^* \tag{4.14}$$

$$\nabla A_j(\bar{z}) v^* = 0 \quad \forall j : \bar{\lambda}_j > 0 \tag{4.15}$$

$$w_j^* = 0 \quad \forall j : \bar{\lambda}_j = 0,\ \nabla A_j(\bar{z}) v^* < 0 \tag{4.16}$$

$$w_j^* \geq 0 \quad \forall j : \bar{\lambda}_j = 0,\ \nabla A_j(\bar{z}) v^* > 0 \tag{4.17}$$

$$F(\bar{x}, \bar{z}) = -\nabla^T A(\bar{z}) \bar{\lambda} \tag{4.18}$$

Proof. The proof is similar to that of Corollary 4.1, only that now, we are forced to use a different transformation formula. Given assumption 3., there exists $v^* \in \mathbb{R}^t$ such that (4.5) and (4.6) hold at $(\bar{x}, \bar{z})$. Clearly (4.5) and (4.13) are the same condition. Moreover, 1. and 2. allow us to apply Corollary 6.2, which states

$$D^* N_C(\bar{z}, -F(\bar{x}, \bar{z}))(v^*) \subseteq \bigcup_{\substack{\bar{\lambda}: \\ F(\bar{x},\bar{z}) = -\nabla^T A(\bar{z})\bar{\lambda}}} \left[\left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z}) v^*) \right],$$

Now, since $D^* N_C \neq \emptyset$, there must exist at least one multiplier $\bar{\lambda} \in \mathbb{R}^p_+$ defined via $F(\bar{x}, \bar{z}) = -\nabla^T A(\bar{z})\bar{\lambda}$ such that $D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z}) v^*) \neq \emptyset$. Thus, by following the same arguments as used in the proof of Corollary 4.1, we obtain (4.14)-(4.18) for each such multiplier $\bar{\lambda}$. □

At this point in time, it is unclear if conditions similar to those found in Corollary 4.1 and 4.2 are attainable when CRCQ[2] is dropped. This is mainly caused by the lack of an explicit transformation formula. See Henrion et al. [2009c, Section 3.4] or in this thesis, Section 6.3 for an indepth discussion on this situation.

4.3 CM-Stationarity

As seen in the proof Corollary 4.2, in cases where C is not a polyhedron, sometimes only an upper approximation of $D^* N_C$ is available. In this section, we define a new concept of stationarity encompassing solutions to the non-linear setting for which only the coderivative to the normal cone mapping to a polyhedral set is required[3].

[2] Actually, CRCQ can be weakened to the need for the calmness of a certain class of multifunctions. See Chapter 6 for more.

[3] By referring to Chapter 6, we see that this can indeed be calculated exactly, regardless of the regularity of the polyhedron

This new stationarity concept, defined via the the next theorem, yields a sharper or at least equivalent type of stationarity conditions to M-stationarity conditions. It should be duly noted that this is the formalization of an idea outlined in Outrata and Červinka [2009] for the nonlinear setting (see Outrata and Červinka [2009, Theorem 2.6] and the discussion following it). However, it is uncertain whether or not these conditions can be obtained under weaker assumptions as the required constraint qualifications allow us to explicitely calculate the contingent cone $T_{\text{gph}\,S}(\bar{x},\bar{z})$, whereas in their absence one may not be able to do so.

Theorem 4.2 (**Critical-M-stationarity of solutions**). *Let $(\bar{x},\bar{z})$ be a local solution to (4.1) and assume that the following conditions hold*

1. *MFCQ holds at $\bar{z} \in C$*
2. *CRCQ holds at $\bar{z} \in C$*
3. *SSOSC holds at $(\bar{x},\bar{z})$*

Then there exist multipliers $v^* \in \mathbb{R}^t$, such that

$$0 \in \nabla_x f(\bar{x},\bar{z}) + \nabla_x^T F(\bar{x},\bar{z})v^* + N_X(\bar{x}) \tag{4.19}$$
$$0 \in \nabla_z f(\bar{x},\bar{z}) + \nabla_z^T F(\bar{x},\bar{z})v^* \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z})\right) v^* + D^* N_{K(\bar{x},\bar{z})}(0,0)(v^*). \tag{4.20}$$

Here, $\bar{\lambda} \in \mathbb{R}^p_+$ is any *Lagrange multiplier associated with the mapping A arising from the relation $F(\bar{x},\bar{z}) = -\nabla^T A(\bar{z})\bar{\lambda}$ and $K(\bar{x},\bar{z}) = T_C(\bar{z}) \cap \{F(\bar{x},\bar{z})\}^\perp$ is the critical cone associated with C corresponding to $(\bar{z},F(\bar{x},\bar{z}))$.*

Proof. Begin by denoting the feasible set to (4.1)

$$\Lambda := \text{gph}\,S \cap \left(X \times \mathbb{R}^t\right).$$

Then since $(\bar{x},\bar{z})$ is a locally optimal solution to (4.1), f is smooth and Λ is closed, we can write the classical first-order optimality conditions for the problem of minimizing a smooth function over a closed set (cf. Rockafellar and Wets [1998, Theorem 6.12]):

$$\langle \nabla_x f(\bar{x},\bar{z}), d\rangle + \langle \nabla_z f(\bar{x},\bar{z}), v\rangle \geq 0, \quad \forall (d,v) \in T_\Lambda(\bar{x},\bar{z}).$$

By virtue of the assumptions, we may use Ralph and Dempe [1995, Theorem 2 and Corollary 4 (statement 2)], which indicate that for any Lagrange multiplier $\bar{\lambda}$ associated with the constraint mapping A at $\bar{z}$, the solution mapping S is single-valued, locally Lipschitz, and directionally differentiable at $\bar{x}$ in direction d, where we denote the directional derivative by $S'(\bar{x};d)$[4]. Moreover, the latter

[4]Though the results in Ralph and Dempe [1995] pertain to the KKT system of a parametric optimization problem, it was later argued in Luo et al. [1997, Theorem 4.2.5 and Lemma 4.2.36 (i)] that these results still hold form generalized equations in the for considered here.

result indicates that $S'(\bar{x};d) = v$, where v is the unique solution of the following generalized equation

$$0 \in \nabla_x F(\bar{x},\bar{z})d + \nabla_z \mathcal{L}(\bar{x},\bar{z},\bar{\lambda})v + N_{K(\bar{x},\bar{z})}(v),$$

regardless of which multiplier $\bar{\lambda}$ is used. Based on the local Lipschitz continuity of S, it is easy to show that

$$T_\Lambda(\bar{x},\bar{z}) = \left\{(d,v) \in T_X(\bar{x}) \times \mathbb{R}^t \,|v = S'(\bar{x};d)\right\}.$$

Then clearly $(0,0)$ is a solution to the following 'linearized' MPEC

$$\min_{d,v} \{\langle \nabla_x f(\bar{x},\bar{z}), d\rangle + \langle \nabla_z f(\bar{x},\bar{z}), v\rangle \mid$$
$$0 \in \nabla_x F(\bar{x},\bar{z})d + \nabla_z \mathcal{L}(\bar{x},\bar{z},\bar{\lambda})v + N_{K(\bar{x},\bar{z})}(v),\ d \in T_X(\bar{x})\right\}. \quad (4.21)$$

Using assumption 3. and Outrata et al. [1998, Theorem 4.6], we observe that the following multifunction arising from the partial linearization with respect to the 'state' variable v of the generalized equation in (4.21), defined

$$\xi \mapsto \left\{v \in \mathbb{R}^t \,\middle|\, \xi \in \nabla_z \mathcal{L}(\bar{x},\bar{z},\bar{\lambda})v + N_{K(\bar{x},\bar{z})}(v)\right\},$$

is single-valued. Moreover, upon observing that $K(\bar{x},\bar{z})$ is polyhedral, Outrata et al. [1998, Corollary 2.5] implies this multifunction is Lipschitz continuous as well and thus by definition, the generalized equation from which the multifunction was defined is strongly regular at $(0,0)$. Therefore, we use Outrata [2000, Proposition 3.2], which indicates that the following constraint qualification holds

$$\left.\begin{array}{c}\begin{bmatrix} 0 & -\nabla_x^T F(\bar{x},\bar{z}) \\ I & -\nabla_z^T \mathcal{L}(\bar{x},\bar{z},\bar{\lambda}) \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} \in N_{T_X(\bar{x})}(0) \times \{0\} \\ (u,w) \in N_{\operatorname{gph} N_{K(\bar{x},\bar{z})}}(0,0) \end{array}\right\} \Rightarrow \begin{array}{c} u = 0 \\ w = 0. \end{array}$$

Then by Outrata [2000, Theorem 3.1], there exist multipliers $v^* \in \mathbb{R}^t$ such that

$$0 \in \nabla_x f(\bar{x},\bar{z}) + \nabla_x^T F(\bar{x},\bar{z})v^* + N_{T_X(\bar{x})}(0)$$
$$0 \in \nabla_z f(\bar{x},\bar{z}) + \nabla_z^T \mathcal{L}(\bar{x},\bar{z},\bar{\lambda})v^* + D^* N_{K(\bar{x},\bar{z})}(0,0)(v^*).$$

Finally, we may use Rockafellar and Wets [1998, Theorem 6.27], from which we know $N_{T_X(\bar{x})}(0) \subset N_X(\bar{x})$. Recalling that the Lagrangian $\mathcal{L}(x,z,\lambda) := F(x,z) + \nabla^T A(z)\lambda$, we see that (4.19) and (4.20) follow. □

We will refer to points for which there exist multipliers $v^* \in \mathbb{R}^t$ such that (4.19) and (4.20) hold as *Critical-Mordukhovich* or simply *CM-stationary* points.

Remark 4.1 (B-stationarity). *The reader shall note that the primal first-order optimality conditions used in the previous proof are essentially what were referred to in this chapter's introduction as B-stationarity conditions. As noted, conditions of this type were first used in terms of MPECs in Luo et al. [1996]. The* *main difficulty in their application has always been in the characterization of the involved contingent cone. Later studies involving this conditions can also be found in Pang and Fukushima [1999], in which the tangent cone was characterized in certain situations, and Scheel and Scholtes [2000], in which connections between strong and B-stationarity were demonstrated.*

Suppose again that $(\bar{x}, \bar{z})$ is a solution to (4.1) and recall that LICQ implies both MFCQ and CRCQ. Then if LICQ holds at $\bar{z}$ and in addition, SSOSC holds at $(\bar{x}, \bar{z})$, the statement of Theorem 4.2 remains valid. Furthermore, given LICQ and SSOSC, Corollary 5.1 indicates via Outrata [2000, Theorem 3.1] that there exists $v^* \in \mathbb{R}^t$ such that (4.5) and (4.6) hold. Thus, under these assumptions, the question arises: "How do CM-stationary solutions relate to M-stationary solutions?", for which we provide the next result.

Proposition 4.1 (**CM-stationarity $\equiv$ M-stationarity**). *Let $(\bar{x}, \bar{z})$ be a local solution to (4.1) and assume the following*

1. *LICQ holds at $\bar{z} \in C$*

2. *SSOSC holds at $(\bar{x}, \bar{z})$*

Then CM-stationarity and M-stationarity are equivalent.

Proof. By the previous discussion, we know there exist multipliers such that both (4.5),(4.6) and (4.19),(4.20) hold at $(\bar{x}, \bar{z})$. Comparing these two sets of relations, it becomes evident that (4.5) and (4.19) are the same condition. Moreover, due to assumption 1., we can rewrite the coderivative in (4.6) with the help of Theorem 6.2

$$D^* N_C(\bar{z}, -F(\bar{x}, \bar{z}))(v^*) = \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z}) v^*),$$

where $\bar{\lambda} \in \mathbb{R}^p_+$ is uniquely defined via $F(\bar{x}, \bar{z})) = -\nabla^T A(\bar{z})\bar{\lambda}$. Upon substituting this formula into (4.6) and comparing to (4.20), we observe that if

$$D^* N_{K(\bar{x},\bar{z})}(0,0)(v^*) = \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z}) v^*),$$

then (4.6) and (4.20) are the same condition. Example 6.1, demonstrates that this is exactly the case. Hence, the assertion holds. $\square$

This last proposition shows us that there is nothing new to be gained by using CM-stationarity conditions in the presence of LICQ and SSOSC. However, CM-stationarity conditions hold under the weaker conditions of MFCQ, CRCQ, and SSOSC. Then as noted, one only needs to calculate the coderivative to a polyhedral set using *any* Lagrange multiplier $\bar{\lambda}$. This stands in stark contrast to the conditions that would be obtained if one knew (4.5) and (4.6) held as well. Indeed, Corollary 6.2 indicates that one can replace the coderivative in (4.6) with an upper-approximation comprised of the union over *all* Lagrange multipliers $\bar{\lambda}$ associated with the mapping A.

4.4 Calculating the Fréchet Normal Cone via Strong Regularity

In this section, we provide results detailing how one could possibly calculate, or at the very least, characterize, a large[5] subset contained within $\widehat{N}_{\mathrm{gph}\, S}$ using strong regularity. These results will then be used later to intimate how the gap between M-stationarity and S-stationarity might be bridged. In the context of MPECs (4.1), results of this type, i.e., using M-stationarity to obtain S-stationarity can also be found in Flegel et al. [2007]. However, though the conditions provided in Flegel et al. [2007] are very weak, their verification may be impossible or require a significant amount of work. In an attempt to partially remedy this problem, we provide an analytical constraint qualification to be used in conjunction with some of the standard constraint qualifications already in use in this chapter. To begin, we associate the solution mapping S with the following generalized equation

$$0 \in \mathcal{F}(x, w) + N_{\mathcal{C}}(w), \tag{4.22}$$

where $x \in \mathbb{R}^s$, $w \in \mathbb{R}^r$, $\mathcal{F} : \mathbb{R}^s \times \mathbb{R}^r \to \mathbb{R}^r$ is continuously differentiable, and $\mathcal{C} \subseteq \mathbb{R}^r$ is a polyhedron.

Proposition 4.2 (polyhedral feasible sets). *Let $(\bar{x}, \bar{w}) \in \mathrm{gph}\, S$ and assume that (4.22) is strongly regular at $(\bar{x}, \bar{w})$. Then*

$$\widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{w}) \supseteq \left\{ \begin{bmatrix} -\nabla_x^T \mathcal{F}(\bar{x}, \bar{w}) v^* \\ u^* - \nabla_w^T \mathcal{F}(\bar{x}, \bar{w}) v^* \end{bmatrix} \middle| u^* \in K^-(\bar{x}, \bar{w}), v^* \in K(\bar{x}, \bar{w}) \right\}. \tag{4.23}$$

Here, $K(\bar{x}, \bar{w}) = T_{\mathcal{C}}(\bar{w}) \cap \{\mathcal{F}(\bar{x}, \bar{w})\}^{\perp}$ is the critical cone to $\mathcal{C}$ corresponding to $(\bar{w}, \mathcal{F}(\bar{x}, \bar{w}))$. Moreover if either of the following conditions holds

1. *$\nabla_x \mathcal{F}(\bar{x}, \bar{w})$ is surjective*
2. *$-\nabla_w \mathcal{F}(\bar{x}, \bar{w}) K(\bar{x}, \bar{w}) \subseteq \mathrm{Im}\, \nabla_x \mathcal{F}(\bar{x}, \bar{w})$*

then the inclusion (4.23) holds as an equality.

Proof. For readability, we leave off the arguments for the critical cone. As $\mathcal{C}$ is a polyhedron, the strong regularity assumption implies that the Lipschitz localization of S, denoted by $\sigma(x)$, is directionally differentiable at $\bar{x}$ for each $d \in \mathbb{R}^s$. Moreover, one has $\sigma'(\bar{x}; d) = v$, where v is the unique solution of the generalized equation

$$0 \in \nabla_x \mathcal{F}(\bar{x}, \bar{w}) d + \nabla_w \mathcal{F}(\bar{x}, \bar{w}) v + N_K(v)$$

(see e.g., Outrata et al. [1998, Theorem 6.3]). We now define

$$\Phi(d, v) := \begin{bmatrix} v \\ -\nabla_x \mathcal{F}(\bar{x}, \bar{w}) d - \nabla_w \mathcal{F}(\bar{x}, \bar{w}) v \end{bmatrix}, \quad \Gamma := \mathrm{gph}\, N_K,$$

[5]Large in the sense that in certain cases, there is significant overlap with this (closed) subset and the limiting normal cone $N_{\mathrm{gph}\, S}$, which is theoretically larger than $\widehat{N}_{\mathrm{gph}\, S}$.

and calculate first the contingent cone to $\operatorname{gph} S$:

$$\begin{aligned}T_{\operatorname{gph} S}(\bar{x},\bar{w}) &= \{(d,v)\in\mathbb{R}^s\times\mathbb{R}^r \,|\, \exists \tau_i \searrow 0, (d_i,v_i)\to(d,v) : \\ &\qquad\qquad \forall i, \bar{w}+\tau_i v_i = \sigma(\bar{x}+\tau_i d_i)\} \\ &= \{(d,v)\in\mathbb{R}^s\times\mathbb{R}^r \,|\, v=\sigma'(\bar{x};d)\}\,,\end{aligned}$$

where the last equality follows from the Lipschitz continuity of σ. Hence,

$$\begin{aligned}T_{\operatorname{gph} S}(\bar{x},\bar{w}) &= \{(d,v)\in\mathbb{R}^s\times\mathbb{R}^r \,| \\ &\qquad 0\in\nabla_x\mathcal{F}(\bar{x},\bar{w})d+\nabla_w\mathcal{F}(\bar{x},\bar{w})v+N_K(v)\} \\ &= \left\{(d,v)\in\mathbb{R}^s\times\mathbb{R}^r \,\middle|\, \begin{bmatrix} v \\ -\nabla_x\mathcal{F}(\bar{x},\bar{w})d-\nabla_w\mathcal{F}(\bar{x},\bar{w})v\end{bmatrix}\in\operatorname{gph} N_K\right\} \\ &= \Phi^{-1}(\Gamma).\end{aligned}$$

Then by definition, $\widehat{N}_{\operatorname{gph} S}(\bar{x},\bar{w}) = [\Phi^{-1}(\Gamma)]^-$. Moreover, given K is a convex cone, it is easy to see that

$$\Gamma = \left\{(v,u)\in K\times K^- \,|\, \langle v,u\rangle = 0\right\}.$$

Clearly, $\Phi^{-1}(K\times K^-)\supset\Phi^{-1}(\Gamma)$. Consequently, by the linearity of Φ,

$$[\Phi^{-1}(\Gamma)]^-\supset[\Phi^{-1}(K\times K^-)]^- = \widehat{N}_{\Phi^{-1}(K\times K^-)}(0,0). \tag{4.24}$$

We claim that

$$\widehat{N}_{\Phi^{-1}(K_1\times K_2)}(0,0) = \nabla^T\Phi(0,0)(K_1^-\times K_2^-) \tag{4.25}$$

for arbitrary polyhedral cones $K_1, K_2\subseteq\mathbb{R}^r$. Indeed, Rockafellar and Wets [1998, Theorem 6.14] implies

$$\widehat{N}_{\Phi^{-1}(K_1\times K_2)}(0,0)\supset\nabla^T\Phi(0,0)\widehat{N}_{K_1\times K_2}(\Phi(0,0)) = \nabla^T\Phi(0,0)(K_1^-\times K_2^-).$$

On the other hand, the multifunction

$$M(p) := \{(a,b)\,|\,\Phi\,(a,b)+p\in K_1\times K_2\}$$

is calm at $(0,0,0)$ due to the polyhedrality of K_1, K_2 and linearity of Φ (cf. Robinson [1981, Proposition 1]). It follows that we can invoke Henrion et al. [2002, Theorem 4.1], which yields the inclusion

$$\begin{aligned}\widehat{N}_{\Phi^{-1}(K_1\times K_2)}(0,0) = N_{\Phi^{-1}(K_1\times K_2)}(0,0) \subset& \\ \nabla^T\Phi(0,0)N_{K_1\times K_2}(\Phi(0,0)) &= \nabla^T\Phi(0,0)(K_1^-\times K_2^-),\end{aligned}$$

whence (4.25). Referring back to (4.24) and given

$$\nabla^T \Phi(0,0) = \begin{bmatrix} 0 & -\nabla_x^T \mathcal{F}(\bar{x},\bar{w}) \\ I & -\nabla_w^T \mathcal{F}(\bar{x},\bar{w}) \end{bmatrix},$$

letting $K_1 = K$ and $K_2 = K^-$ proves (4.23) holds.

Now note that if we can demonstrate that the reverse inclusion of (4.24) holds, then (4.23) holds as an equality. Assume first that $\nabla_x \mathcal{F}(\bar{x},\bar{w})$ is surjective. We observe that both sets $K \times \{0\}$ and $\{0\} \times K^-$ are subsets of Γ. Furthermore, by taking into account (4.25) with appropriate settings for K_1 and K_2, one has

$$\begin{aligned} [\Phi^{-1}(\Gamma)]^- &\subset [\Phi^{-1}(K \times \{0\})]^- \cap [\Phi^{-1}(\{0\} \times K^-)]^- \\ &= \widehat{N}_{\Phi^{-1}(K\times\{0\})}(0,0) \cap \widehat{N}_{\Phi^{-1}(\{0\}\times K^-)}(0,0) \\ &= \nabla^T\Phi(0,0)(K^- \times \mathbb{R}^r) \cap \nabla^T\Phi(0,0)(\mathbb{R}^r \times K) \\ &= \nabla^T\Phi(0,0)(K^- \times K) \end{aligned}$$

where the last equality holds since $\nabla\Phi(0,0)$ is surjective. Thus, (4.24) and accordingly, (4.23), holds as an equality.

To see that the second condition in the theorem statement also implies (4.23) holds as an equality, begin by noting that

$$\Phi^{-1}(\Gamma) = \left\{ (d,v) \in \mathbb{R}^s \times \mathbb{R}^r \;\middle|\; \begin{array}{c} v \in K \\ -\nabla_x \mathcal{F}(\bar{x},\bar{w})d - \nabla_w \mathcal{F}(\bar{x},\bar{w})v \in K^- \\ \langle v, \nabla_x \mathcal{F}(\bar{x},\bar{w})d + \nabla_w \mathcal{F}(\bar{x},\bar{w})v \rangle = 0 \end{array} \right\}$$

Let $(d^*, v^*) \in [\Phi^{-1}(\Gamma)]^-$ be arbitrary, then it holds that

$$\begin{gathered} \langle d^*, d\rangle + \langle v^*, v\rangle \le 0 \\ \forall v \in K,\, \forall d : \begin{array}{c} -\nabla_x \mathcal{F}(\bar{x},\bar{w})d - \nabla_w \mathcal{F}(\bar{x},\bar{w})v \in K^- \\ \langle v, \nabla_x \mathcal{F}(\bar{x},\bar{w})d + \nabla_w \mathcal{F}(\bar{x},\bar{w})v \rangle = 0. \end{array} \end{gathered} \tag{4.26}$$

As $0 \in K$, setting $v = 0$ implies

$$\langle d^*, d\rangle \le 0, \quad -\nabla_x \mathcal{F}(\bar{x},\bar{w})d \in K^-.$$

Then since

$$\begin{aligned} \left\{ d \,\middle|\, -\nabla_x \mathcal{F}(\bar{x},\bar{w})d \in K^- \right\} &= \{ d \,|\, \langle -\nabla_x \mathcal{F}(\bar{x},\bar{w})d, u\rangle \le 0,\ \forall u \in K \} \\ &= \left\{ d \,\middle|\, \langle d, -\nabla_x^T \mathcal{F}(\bar{x},\bar{w})u\rangle \le 0,\ \forall u \in K \right\} \\ &= \left\{ d \,\middle|\, \langle d, y\rangle \le 0,\ \forall y \in -\nabla_x^T \mathcal{F}(\bar{x},\bar{w})K \right\} \\ &= \left[-\nabla_x^T \mathcal{F}(\bar{x},\bar{w})K \right]^-, \end{aligned}$$

it follows that

$$d^* \in \left[\left[-\nabla_x^T \mathcal{F}(\bar{x},\bar{w})K\right]^-\right]^- = -\nabla_x^T \mathcal{F}(\bar{x},\bar{w})K,$$

as K is a closed convex cone. Consequently, there exists a $\bar{u} \in K$ such that

$$d^* = -\nabla_x^T \mathcal{F}(\bar{x},\bar{w})\bar{u}. \tag{4.27}$$

Relation (4.26) now yields

$$\langle -\nabla_x^T \mathcal{F}(\bar{x},\bar{w})\bar{u}, d\rangle + \langle v^*, v\rangle \leq 0$$

$$\forall v \in K \text{ and } \forall d : \begin{array}{c} -\nabla_x \mathcal{F}(\bar{x},\bar{w})d - \nabla_w \mathcal{F}(\bar{x},\bar{w})v \in K^- \\ \langle v, \nabla_x \mathcal{F}(\bar{x},\bar{w})d + \nabla_w \mathcal{F}(\bar{x},\bar{w})v\rangle = 0, \end{array}$$

that is

$$\langle \bar{u}, -\nabla_x \mathcal{F}(\bar{x},\bar{w})d\rangle + \langle v^*, v\rangle \leq 0$$

$$\forall v \in K \text{ and } \forall d : \begin{array}{c} -\nabla_x \mathcal{F}(\bar{x},\bar{w})d - \nabla_w \mathcal{F}(\bar{x},\bar{w})v \in K^- \\ \langle v, \nabla_x \mathcal{F}(\bar{x},\bar{w})d + \nabla_w \mathcal{F}(\bar{x},\bar{w})v\rangle = 0. \end{array}$$

Now let v be arbitrary. By assumption 2., there exists a d such that $-\nabla_x \mathcal{F}(\bar{x},\bar{w})d = \nabla_w \mathcal{F}(\bar{x},\bar{w})v$. Accordingly, we can now derive from the previous relation

$$\langle \bar{u}, \nabla_w \mathcal{F}(\bar{x},\bar{w})v\rangle + \langle v^*, v\rangle \leq 0.$$

As v is arbitrary, we have:

$$\langle v^* + \nabla_w^T \mathcal{F}(\bar{x},\bar{w})\bar{u}, v\rangle \leq 0, \ \forall v \in K,$$

in other words, it holds that

$$v^* + \nabla_w^T \mathcal{F}(\bar{x},\bar{w})\bar{u} \in K^-.$$

Then along with (4.27), using $\bar{v} = v^* + \nabla_w^T \mathcal{F}(\bar{x},\bar{w})\bar{u}$, we have

$$(d^*, v^*) = \begin{bmatrix} 0 & -\nabla_x^T \mathcal{F}(\bar{x},\bar{w}) \\ I & -\nabla_w^T \mathcal{F}(\bar{x},\bar{w}) \end{bmatrix} \begin{pmatrix} \bar{v} \\ \bar{u} \end{pmatrix}$$

Since $\bar{v} \in K^-$ and $\bar{u} \in K$ the inclusion (4.24) holds in the opposite direction, thus making (4.23) and equality. $\square$

We now use Proposition 4.2 to obtain a similar statement for the solution mapping S to the generalized equation (4.2), i.e., for settings in which a non-polyhedral convex feasible set is considered. We start by writing the so-called *enhanced generalized equation* associated with (4.2)

$$0 \in \begin{bmatrix} \mathcal{L}(x,z,\lambda) \\ -A(z) \end{bmatrix} + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z,\lambda), \tag{4.28}$$

where, as before,

$$\mathcal{L}(x, z, \lambda) = F(x, z) + \nabla^T A(z)\lambda$$

and λ is a vector of Lagrange multipliers associated with the constraint mapping A. For the enhanced generalized equation, we introduce the *enhanced solution mapping*

$$S^e(x) := \Big\{(z, \lambda) \in \mathbb{R}^t \times \mathbb{R}^p \,|(4.28) \text{ is fulfilled}\Big\}. \tag{4.29}$$

Clearly (4.28) is of the form (4.22), where

$$w := (z, \lambda)\,, \quad \mathcal{F}\,(x, w) := \begin{bmatrix} \mathcal{L}(x, z, \lambda) \\ -A(z) \end{bmatrix}, \quad \mathcal{C} := \mathbb{R}^t \times \mathbb{R}^p_+.$$

On the basis of Proposition 4.2 we arrive now at the following statement.

Proposition 4.3 (an inner approximation of $\widehat{N}_{\operatorname{gph} S^e}$). *Consider a reference point $(\bar{x}, \bar{z}, \bar{\lambda}) \in$ gph S^e and assume that (4.28) is strongly regular at $(\bar{x}, \bar{z}, \bar{\lambda})$. Then,*

$$\begin{aligned}\widehat{N}_{\operatorname{gph} S^e}(\bar{x}, \bar{z}, \bar{\lambda}) \supseteq \{(a, b, c) \in \mathbb{R}^s \times \mathbb{R}^t \times \mathbb{R}^p \mid\ & \exists v \in \mathbb{R}^t, u \in \mathbb{R}^{a_+} \times \mathbb{R}^{a_0}_+ \times \{0\},\\ & u' \in \{0\} \times \mathbb{R}^{a_0}_- \times \mathbb{R}^{p-a_0-a_+} :\\ & a = -\nabla^T_x F(\bar{x}, \bar{z})v\\ & b = -\nabla^T_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda})v + \nabla^T A_{I_+}(\bar{z})u_{I_+} + \nabla^T A_{I_0}(\bar{z})u_{I_0}\\ & c_{I_+} = -\nabla A_{I_+}(\bar{z})v\\ & c_{I_0} = u'_{I_0} - \nabla A_{I_0}(\bar{z})v\\ & c_L = u'_L - \nabla A_L(\bar{z})v\},\end{aligned}$$

where $\nabla^T_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda})v = [\nabla^T_z F(\bar{x}, \bar{z}) + \sum_{i=1}^p \bar{\lambda}_i \nabla^2 A_i(\bar{z})]v$, $a_+ := |I_+|$, *and* $a_0 := |I_0|$.

Proof. Letting $w = (z, \lambda)$ in Proposition 4.2, it suffices to compute

$$\begin{aligned} K(\bar{x}, \bar{z}, \bar{\lambda}) &= T_{\mathbb{R}^t \times \mathbb{R}^p_+}(\bar{z}, \bar{\lambda}) \cap \begin{bmatrix} \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) \\ -A(\bar{z}) \end{bmatrix}^\perp \\ &= \Big\{(v, u) \in \mathbb{R}^t \times \mathbb{R}^p \,|u_{L \cup I_0} \geq 0\Big\} \cap \begin{bmatrix} 0 \\ -A(\bar{z}) \end{bmatrix}^\perp \\ &= \Big\{(v, u) \in \mathbb{R}^t \times \mathbb{R}^p \,\Big| -A(\bar{z})^T u = 0,\ u_{L \cup I_0} \geq 0\Big\} \\ &= \Big\{(v, u) \in \mathbb{R}^t \times \mathbb{R}^p \,|u_L = 0, u_{I_0} \geq 0\Big\} \end{aligned}$$

and

$$K^-(\bar{x}, \bar{z}, \bar{\lambda}) = \Big\{(v', u') \in \mathbb{R}^m \times \mathbb{R}^p \,\Big| v' = 0, u'_{I_+} = 0, u'_{I_0} \leq 0\Big\}$$

and apply Proposition 4.2 with

$$\begin{aligned}
\nabla_x \mathcal{F}(\bar{x}, \bar{w}) &= \begin{bmatrix} \nabla_x F(\bar{x}, \bar{z}) \\ 0 \end{bmatrix} \\
\nabla_w \mathcal{F}(\bar{x}, \bar{w}) &= \begin{bmatrix} \nabla_z F(\bar{x}, \bar{z}) + \sum_{i=1}^{p} \bar{\lambda}_i \nabla^2 A_i(\bar{z}) & \nabla^T A(\bar{z}) \\ -\nabla A(\bar{z}) & 0 \end{bmatrix}.
\end{aligned}$$

□

Remark 4.2. *Due to the fact that both variables u'_L and v are free, the component c_L in the statement of Proposition 4.3 becomes inconsequential.*

Remark 4.3 (equality in Proposition 4.3). *As shown in Proposition 4.2, if either of the two conditions provided there holds, then (4.23) holds as an equality. First note that in the current setting, condition 1. will never hold. Assume for the sake of argument that condition 2. holds in the context of (4.28). Then for all $(v, u) \in K(\bar{x}, \bar{z}, \bar{\lambda})$ there must exist an $\eta \in \mathbb{R}^s$ such that*

$$\begin{gathered}
\nabla_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) v + \nabla^T A(\bar{z}) u = \nabla_x F(\bar{x}, \bar{z}) \eta \\
-\nabla A(\bar{z}) v = 0
\end{gathered}$$

Such a condition however, would most likely only hold in a select number of cases. However, this does not rule out other settings where $\mathcal{F}$ has a different form.

Based on the structure provided by Proposition 4.3, we next compute a similar inner approximation for $\widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{z})$, where S is the solution map associated with (4.2).

Proposition 4.4 (an inner approximation of $\widehat{N}_{\operatorname{gph} S}$). *Under the assumptions of Proposition 4.3, we have that*

$$\widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{z}) \supseteq \left\{ \begin{bmatrix} \nabla_x^T F(\bar{x}, \bar{z}) v^* \\ \nabla_z^T \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) v^* + \nabla^T A(\bar{z}) w^* \end{bmatrix} \middle| \forall v^*, \forall w^* : \begin{array}{r} \nabla A_{I_+}(\bar{z}) v^* = 0 \\ \nabla A_{I_0}(\bar{z}) v^* \geq 0 \\ w^*_{I_0} \geq 0 \end{array} \right\}. \tag{4.30}$$

Proof. Note that due to strong regularity, $\bar{\lambda}$, used to defined I_+ and I_0, is the unique multiplier vector associated with the pair $(\bar{x}, \bar{z})$. We claim that

$$\widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{z}) = \left\{ (a, b) \in \mathbb{R}^s \times \mathbb{R}^t \,\middle|\, (a, b, 0) \in \widehat{N}_{\operatorname{gph} S^e}(\bar{x}, \bar{z}, \bar{\lambda}) \right\}. \tag{4.31}$$

Indeed, by Rockafellar and Wets [1998, Theorem 6.11], one has $(a, b) \in \widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{z})$ if and only if there is a smooth function h that achieves its local maximum relative to $\operatorname{gph} S$ at $(\bar{x}, \bar{z})$ and $\nabla h(\bar{x}, \bar{z}) = (a, b)$. Then clearly $(\bar{x}, \bar{z}, \bar{\lambda})$ is a local maximum of the function $\widetilde{h}$ on $\operatorname{gph} S^e$, where

$$\widetilde{h}(x, z, \lambda) = h(x, z) \text{ for all } \lambda.$$

Consequently, $(a, b, 0) \in \widehat{N}_{\operatorname{gph} S^e}(\bar{x}, \bar{z}, \bar{\lambda})$. For the reverse direction, we appeal to the equivalent definition of the Fréchet normal cone (see e.g., Mordukhovich [2006a, Definition 1.1]), which states

$$(a, b, 0) \in \widehat{N}_{\operatorname{gph} S^e}(\bar{x}, \bar{z}, \bar{\lambda}) \Leftrightarrow \limsup_{\substack{(x,z,\lambda)\to(\bar{x},\bar{z},\bar{\lambda}) \\ (x,z,\lambda)\in\operatorname{gph} S^e \\ (x,z,\lambda)\neq(\bar{x},\bar{z},\bar{\lambda})}} \frac{\langle a, x-\bar{x}\rangle + \langle b, z-\bar{z}\rangle + 0}{||(x,z,\lambda)-(\bar{x},\bar{z},\bar{\lambda})||} \leq 0.$$

We claim now that this implies $(a, b) \in \widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{z})$. Indeed, due to the strong regularity assumption, both z and λ are single-valued locally Lipschitz functions of x near $\bar{x}$. In particular, it is easy to argue that locally around $(\bar{x}, \bar{z}, \bar{\lambda})$ one has that

$$(x, z) \in \operatorname{gph} S \Leftrightarrow (x, z, \lambda(x)) \in \operatorname{gph} S^e.$$

Then we may continue the inequality given above as follows

$$0 \geq \limsup_{\substack{(x,z)\to(\bar{x},\bar{z}) \\ (x,z)\in\operatorname{gph} S \\ (x,z)\neq(\bar{x},\bar{z})}} \frac{\langle a, x-\bar{x}\rangle + \langle b, z-\bar{z}\rangle}{||(x,z,\lambda(x))-(\bar{x},\bar{z},\bar{\lambda})||} \geq \frac{1}{L+1} \limsup_{\substack{(x,z)\to(\bar{x},\bar{z}) \\ (x,z)\in\operatorname{gph} S \\ (x,z)\neq(\bar{x},\bar{z})}} \frac{\langle a, x-\bar{x}\rangle + \langle b, z-\bar{z}\rangle}{||x-\bar{x}|| + ||z-\bar{z}||}$$

where L is the Lipschitz modulus of $\lambda(x)$. Thus by definition, $(a, b) \in \widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{z})$, which proves (4.31). Then the asserted formula follows immediately from Proposition 4.3. □

In the event that strict complementarity holds, the solution mapping S is smooth. Therefore, by writing the generalized equation (4.2) down in the enhanced form, i.e., via equations and complementarity relations, one could use the classical implicit function theorem to obtain a representation of $\widehat{N}_{\operatorname{gph} S}$. Indeed, in this case, the Fréchet normal cone would amount to $\{\alpha \nabla S(\bar{x}) \,|\, \alpha \geq 0\}$.

4.5 Explicit S-Stationarity Conditions

We can now use the results of the previous section to obtain explicit S-stationarity conditions. First, we use the inclusion (4.30) to define a type of *pseudo stationarity* conditions. That is, we say that a feasible point $(\bar{x}, \bar{z})$ to the MPEC (4.1) is pseudo-stationary, if there exist $\bar\lambda \in \mathbb{R}^p_+$ and $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ such that

$$-\nabla_x f(\bar x, \bar z) = \nabla_x^T F(\bar x, \bar z) v^* \tag{4.32}$$

$$-\nabla_z f(\bar x, \bar z) = \nabla_z^T F(\bar x, \bar z) v^* + \left(\sum_{j=1}^p \bar\lambda_j \nabla^2 A_j(\bar z) \right) v^* + \nabla^T A(\bar z) w^* \tag{4.33}$$

$$\nabla A_{I_+}(\bar z) v^* = 0 \tag{4.34}$$

$$w^*_{I_0} \geq 0, \forall j : \bar\lambda_j = 0, \nabla A_j(\bar z) v^* \geq 0 \tag{4.35}$$

$$F(\bar x, \bar z) = -\nabla^T A(\bar z) \bar\lambda \tag{4.36}$$

Though we have used the righthand side of the inclusion in Proposition 4.4 to define the pseudo-stationarity conditions, these conditions are not *true* stationarity conditions, as a solution to the MPEC must not satisfy them. That is, there do not have to exist (v^*, w^*) and $\bar\lambda$ such that (4.32)-(4.36) hold at a solution. Nevetheless, if we can obtain a tuple $(\bar x, \bar z, \bar\lambda, v^*, w^*)$ such that $(\bar x, \bar z)$ *is* a solution to the MPEC and (4.32)-(4.36) hold, provided strong regularity of the enhanced generalized equation, then these conditions amount to true strong stationarity conditions. In order to do so, we provide a condition ensuring that M-stationary tuples satisfying (4.7)-(4.12) satisfy (4.32)-(4.36) as well.

Theorem 4.3 (**using M-stationarity to obtain S-stationarity**). *Let $(\bar x, \bar z)$ be a local solution to (4.1) and assume without loss of generality that $A(\bar z) = 0$. If the following conditions hold*

1. *$X \equiv \mathbb{R}^s$*
2. *LICQ at $\bar z \in C$*
3. *SSOSC at $(\bar x, \bar z)$*
4. *there exists a $v^* \in \mathbb{R}^t$ such that the triple $(\bar x, \bar z, v^*)$ satisfies (4.5) and (4.6), $\nabla A_{I_0}(\bar z) v^* > 0$.*

Then the the following conditions are S-stationarity conditions, where v^* is as in assumption 4.: there exists a unique $\bar\lambda \in \mathbb{R}^p_+$ and $w^* \in \mathbb{R}^p$ such that

$$-\nabla_x f(\bar x, \bar z) = \nabla_x^T F(\bar x, \bar z) v^*$$

$$-\nabla_z f(\bar x, \bar z) = \nabla_z^T F(\bar x, \bar z) v^* + \left(\sum_{j=1}^p \bar\lambda_j \nabla^2 A_j(\bar z) \right) v^* + \nabla^T A(\bar z) w^*$$

$$\nabla A_{I_+}(\bar z) v^* = 0$$

$$\nabla A_{I_0}(\bar z) v^* > 0$$

$$w^*_{I_0} \geq 0$$

$$F(\bar x, \bar z) = -\nabla^T A(\bar z) \bar\lambda$$

Proof. Given LICQ and SSOSC, we know (4.28) is strongly regular at $(\bar{x}, \bar{z}, \bar{\lambda})$ (see Robinson [1980, Theorem 4.1]). Thus, the inclusion (4.30) holds and we can write the pseudo-stationarity conditions (4.32)-(4.36). Moreover, LICQ and SSOSC imply that the perturbation mapping associated with (4.2) has the Aubin property at $(\bar{x}, \bar{z})$ (see Corollary 5.1), in which case we can apply Corollary 4.1, which states that there exists a unique $\bar{\lambda} \in \mathbb{R}^p_+$ and multipliers $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ such that (4.7)-(4.12) hold. Under assumption 4., we have a v^* such that there exists a unique $\bar{\lambda} \in \mathbb{R}^p_+$ and vectors $w^* \in \mathbb{R}^p$ with

$$\begin{aligned} -\nabla_x f(\bar{x}, \bar{z}) &= \nabla_x^T F(\bar{x}, \bar{z}) v^* \\ -\nabla_z f(\bar{x}, \bar{z}) &= \nabla_z^T F(\bar{x}, \bar{z}) v^* + \left(\sum_{j=1}^p \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) w^* \end{aligned}$$

$$\begin{aligned} \nabla A_{I_+}(\bar{z}) v^* &= 0 \\ \nabla A_{I_0}(\bar{z}) v^* &> 0 \\ w^*_{I_0} &\geq 0 \\ F(\bar{x}, \bar{z}) &= -\nabla^T A(\bar{z}) \bar{\lambda} \end{aligned}$$

Then the tuple $(\bar{x}, \bar{z}, \bar{\lambda}, v^*, w^*)$ satisfies the pseudo-stationarity conditions, which by the previous discussion implies that these are in fact strong stationarity conditions. □

Note that if strict complementarity were to hold, then the pseudo- stationarity conditions would always coincide with the M-stationarity conditions, in which case M-stationarity and S-stationarity become equivalent conditions. We refer the reader to Example 8.3 for a case where the assumptions of Theorem 4.3 hold even though strict complementarity fails.

4.6 Stationarity Conditions for EPECs

Turning our attention now back to EPECs, we form the EPEC of interest by coupling n MPECs of the type (4.1) as described in Chapter 3, letting $x = (x_1, \dots, x_n) \in \mathbb{R}^{ns}$. As we will be considering EPECs in which $X \equiv \mathbb{R}^s$, we leave out such constraints in the following discussion and consider the EPEC

$$\min_{x_i, z} \{ f_i(x_{-i}, x_i, z) \, | \, z \in S(x_{-i}, x_i) \} \quad (i = 1, \dots, n). \tag{4.37}$$

Recall the notation $S_i(x_i) = S_{x_{-i}}(x_i) = S(x_{-i}, x_i)$ and consider the following definition.

Definition 4.3 (**M- and S-stationarity conditions for EPECs**). *A feasible strategy $(\hat{x}, \hat{z})$ to (4.37) is called* ***strongly*** *or* ***S-stationary*** *if for all $i = 1, \dots, n$, the following condition holds*

$$0 \in \nabla_{x_i, z} f(\hat{x}_{-i}, \hat{x}_i, \hat{z}) + \widehat{N}_{\operatorname{gph} S_i}(\hat{x}_i, \hat{z}). \tag{4.38}$$

Similarly, a feasible strategy $(\hat{x}, \hat{z})$ to (4.37) is called ***Mordukhovich*** *or* ***M-stationary*** *if for all $i = 1, \ldots, n$, the following condition holds*

$$0 \in \nabla_{x_i,z} f(\hat{x}_{-i}, \hat{x}_i, \hat{z}) + N_{\operatorname{gph} S_i}(\hat{x}_i, \hat{z}). \tag{4.39}$$

Notice how (4.38) and (4.39) are composed of the S- and M-stationarity conditions for the MPECs making up the EPEC (4.37), respectively. This makes the application of many of the results from the previous sections rather simple. Indeed, the key to seeing that the majority of results presented in the previous sections can be translated into the EPEC setting, lies in the observation that if the solution mapping S or the perturbation mapping Ψ have certain stability properties, then these properties will carry over to the more "restrictive" settings. For example, if the generalized equation

$$0 \in F(x, z) + N_C(z)$$

is strongly regular at $(\bar{x}, \bar{z})$, then the more restricted generalized equation

$$0 \in F(\bar{x}_{-i}, x, z) + N_C(z)$$

is strongly regular at $(\bar{x}_i, \bar{z})$ as well. Indeed, strong regularity at $(\bar{x}, \bar{z})$ implies that the solution mapping of the partial linearization

$$\eta \in F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) N_C(z)$$

has a single-valued Lipschitz localization near $(0, \bar{z})$. Then since the solution mapping of the partial linearization of the restricted generalized equation at $(\bar{x}_i, \bar{z})$ is the same as the previous one, it too has a single-valued Lipschitz localization near $(0, \bar{z})$. Also note that in our setting, LICQ, MFCQ, CRCQ and SSOSC are not affected by the coupling process used to define EPECs.

This concludes our discussion of stationarity conditions for MPECs and EPECs. For a more indepth discussion of EPEC stationarity conditions, the reader is referred to the recent theses by Červinka [2008] and Su [2005].

Chapter 5

Stability of the Perturbation Mapping

Based on the results provided in Chapter 4, we observe that in order to use the stationarity conditions, we need to know which stability properties the solution mappings S and perturbation mappings Ψ possess. Upon referencing Corollaries 4.1 and 4.2, we see that the fundamental stability property of interest is calmness. Fortunately, there exists a wide range of results in the literature providing criteria for checking the presence of various stability properties of multifunctions, of which there appears to be two main approaches: those utilizing generalized derivatives and those which do not. Conditions using generalized derivatives based on contingent cones were first developed in Aubin [1981] and further developed and applied to settings similar to ours in Klatte and Kummer [1999, 2002b, 2003]. For the application of generalized derivatives using the limiting variational objects we direct the reader to Mordukhovich [1993, 2006a], for checking the Aubin property, Henrion et al. [2002], for calmness using coderivative conditions, and recently, Ioffe and Outrata [2008], using newly developed generalized-derivative-like objects for calmness. For approaches to stability without the usage of generalized derivatives, using instead specially developed algorithms, we refer the reader to the more recent works Heerda and Kummer [2006], Klatte and Kummer [2009]. This however, should *not* be taken as an exhaustive list. Thus, the main purpose of this chapter is to provide a few supplementary results seeking to fill some of the gaps in the literature, thereby facilitating a more encompassing analysis of our example problem in the latter chapters of this thesis.

5.1 Calmness of Perturbation Mappings via Strong Regularity

Our framework in this section is as follows. We begin by letting S be the solution mapping of the generalized equation

$$0 \in F(x, z) + N_C(z), \tag{5.1}$$

where $F : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}^t$ is continuously differentiable and the feasible set $C := \{z \in \mathbb{R}^t \,|\, A_j(z) \leq 0, j = 1, \ldots, p\}$ with $A_j(z)$ twice continuously differentiable and convex for all $j = 1, \ldots, p$ and define the associated perturbation mapping

$$\Psi(u) := \{(x,z)\,|u \in F(x,z) + N_C(z)\}\,.$$

In Outrata [2000, Proposition 3.2], it was demonstrated that if (5.1) is strongly regular at $(\bar{x}, \bar{z}) \in \operatorname{gph} S$ in the sense Robinson, then the multifunction

$$\widehat{\Psi}(u_1, u_2) := \{(x,z)\,|u_2 \in F(x,z) + N_C(z + u_1)\}$$

has the Aubin property at $(0, 0, \bar{x}, \bar{z})$ and is therefore calm as well. Clearly this implies the Aubin property for the perturbation mapping Ψ at $(0, \bar{x}, \bar{z})$.

Recall from Robinson [1980, Theorem 4.1] that the enhanced generalized equation of (5.1), i.e.,

$$0 \in \begin{bmatrix} \mathcal{L}(x,z,\lambda) \\ -A(z) \end{bmatrix} + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda) \tag{5.2}$$

is strongly regular at $(\bar{x}, \bar{z}, \bar{\lambda})$ provided LICQ holds at $\bar{z}$ and SSOSC at $(\bar{x}, \bar{z})$. And though it is easy to show that this implies S is a single-valued locally Lipschitz function near $(\bar{x}, \bar{z}) \in \operatorname{gph} S$, it does not directly indicate that the solution mapping of the partial linearization of (5.1), i.e.,

$$\xi \in F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) + N_C(z), \tag{5.3}$$

is single-valued and locally Lipschitz near $(0, \bar{z})$ with $\bar{\xi} = 0$ when C is non-polyhedral[1], that is, (5.1) is strongly regular in the sense of Robinson. Therefore, we provide the following result, which to our knowledge has not been formulated as it is here in the literature and which is integral in building certain arguments of this thesis.

Proposition 5.1. *Let $(\bar{x}, \bar{z})$ be a solution to (5.1) and assume that*

1. *LICQ holds at $\bar{z}$*
2. *SSOSC holds at $(\bar{x}, \bar{z})$*

Then (5.1) is strongly regular in the sense of Robinson at $(\bar{x}, \bar{z})$.

Proof. As mentioned in the preceeding discussion, the assumptions imply via Robinson [1980, Theorem 4.1] that (5.2) is strongly regular at $(\bar{x}, \bar{z}, \bar{\lambda})$. Consequently, the solution mapping arising from the partial linearization of (5.2) defined

$$\Sigma(\eta) := \left\{(z,\lambda)\,\middle|\,\eta \in C(\bar{x}, \bar{z}, \bar{\lambda}) + \nabla_{z,\lambda} C(\bar{x}, \bar{z}, \bar{\lambda})((z,\lambda) - (\bar{z}, \bar{\lambda})) + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda)\right\},$$

where

$$C(x,z,\lambda) := \begin{pmatrix} \mathcal{L}(x,z,\lambda) \\ -A(z) \end{pmatrix} = \begin{pmatrix} F(x,z) + \nabla^T A(z)\lambda \\ -A(z) \end{pmatrix},$$

[1] In the event C is polyhedral, Dontchev and Rockafellar [1996, Theorem 3] implies that these properties are equivalent.

is single-valued and locally Lipschitz near $(0, \bar{z}, \bar{\lambda})$. Next, we define

$$\Phi(\xi, z, \lambda) := \begin{pmatrix} F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) + \nabla^T A(z)\lambda \\ -A(z) \end{pmatrix} - \xi,$$

from which we can immediately see that

$$C(\bar{x}, \bar{z}, \bar{\lambda}) = \Phi(0, \bar{z}, \bar{\lambda}) \quad \nabla_{x,\lambda} C(\bar{x}, \bar{z}, \bar{\lambda}) = \nabla_{z,\lambda}\Phi(0, \bar{z}, \bar{\lambda}).$$

Consequently,

$$\Sigma(\eta) = \left\{(z, \lambda) \,\middle|\, \eta \in \Phi(0, \bar{z}, \bar{\lambda}) + \nabla_{z,\lambda}\Phi(0, \bar{z}, \bar{\lambda})((z, \lambda) - (\bar{z}, \bar{\lambda})) + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda)\right\}.$$

Since we already know Σ to be a single-valued locally Lipschitz function around $(0, \bar{z}, \bar{\lambda})$, it follows that the generalized equation

$$0 \in \Phi(\xi, z, \lambda) + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda)$$

is strongly regular at $(0, \bar{z}, \bar{\lambda})$ and hence, the mapping

$$\xi \mapsto \left\{(z, \lambda) \,\middle|\, 0 \in \Phi(\xi, z, \lambda) + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda)\right\}$$

is single-valued and locally Lipschitz near $(0, \bar{z}, \bar{\lambda})$. This then amounts to saying that the mapping

$$\xi \mapsto \left\{(z, \lambda) \,\middle|\, \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in \begin{pmatrix} F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) + \nabla^T A(z)\lambda \\ -A(z) \end{pmatrix} + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda)\right\} \tag{5.4}$$

is single-valued and locally Lipschitz near $(0, \bar{z}, \bar{\lambda})$ and therefore, so is the following mapping as well

$$\xi \mapsto \left\{(z, \lambda) \,\middle|\, \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} \in \begin{pmatrix} F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) + \nabla^T A(z)\lambda \\ -A(z) \end{pmatrix} + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda)\right\}$$

We now claim that this also implies the solution mapping to (5.3), i.e.,

$$\xi \mapsto \{z \,|\, \xi \in F(\bar{x}, \bar{z}) + \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) + N_C(z)\}, \tag{5.5}$$

has a single-valued Lipschitz localization at $(0, \bar{z})$. Indeed, begin by noting that due to LICQ, i.e., $\nabla A(\bar{z})$ is surjective,

$$N_C(z) = \left\{ v \,\middle|\, \exists \lambda : v = \nabla^T A(z)\lambda, A(z) \in N_{\mathbb{R}^p_+}(\lambda) \right\}$$

for all z close to $\bar{z}$, since surjectivity is a local property. Moreover, using again the surjectivity of $\nabla A(\bar{z})$ and the fact that $-F(\bar{x}, \bar{z}) = \nabla^T A(\bar{z})\bar{\lambda}$, the equation

$$\xi_1 - F(\bar{x}, \bar{z}) - \nabla_z F(\bar{x}, \bar{z})(z - \bar{z}) = \nabla^T A(z)\lambda$$

has a unique solution λ which is close to $\bar{\lambda}$ if (ξ_1, z) is close to $(0, \bar{z})$. These two observations allow to transfer the local Lipschitz and uniqueness statement from (5.4) to (5.5). This however, means that the generalized equations

$$0 \in F(x, z) + N_C(z)$$

are strongly regular at $(\bar{x}, \bar{z})$. □

Using Proposition 5.1, we have the following implication.

Corollary 5.1. *Let $(\bar{x}, \bar{z})$ be a solution to (5.1). If the following assumptions hold*

1. *LICQ is satisfied at $\bar{z}$*
2. *SSOSC is satisfied at $(\bar{x}, \bar{z})$*

Then Ψ has the Aubin property at $(0, \bar{x}, \bar{z})$.

Proof. Given the assumptions, Proposition 5.1 implies that (5.1) is strongly regular at $(\bar{x}, \bar{z})$. Thus, referring to Outrata [2000, Proposition 3.2], we know that the mapping

$$(u_1, u_2) \mapsto \{(x, z) \,|u_2 \in F(x, z) + N_C(u_1 + z)\}$$

has the Aubin property at $(0, 0, \bar{x}, \bar{z})$. Then clearly, the more the restricted mapping

$$(0, u_2) \mapsto \{(x, z) \,|u_2 \in F(x, z) + N_C(z)\},$$

which amounts to Ψ, has the Aubin property at $(0, \bar{x}, \bar{z})$ as well. □

Corollary 5.1 demonstrates that Ψ has, in this case, the stronger Aubin property and is therefore calm at $(0, \bar{x}, \bar{z})$. As we will see in Example 7.1, when SSOSC does not hold, it is not always the case that Ψ has the Aubin property even for a very simple example. Therefore, in order to counter this problem, we provide a general result in the next section.

In many of the recent papers utilizing calmness as a constraint qualification for certain calculus rules of limiting variational objects, e.g., Henrion et al. [2002], one requires not calmness of the perturbation mapping, but instead that of the mapping $\widehat{\Psi}$ defined at the beginning of this section. However, the following result shows the the calmness of Ψ is equivalent to that of $\widehat{\Psi}$. This was first shown by Outrata [2008], however, as it has yet to be published, we provide the proof as well.

Proposition 5.2. *Let* $(\bar{x}, \bar{z}) \in \operatorname{gph} S$. *Then the perturbation mapping* Ψ *associated with (5.1) is calm at* $(0, \bar{x}, \bar{z})$ *if and only if the mapping* $\widehat{\Psi}$ *is calm at* $(0, 0, \bar{x}, \bar{z})$.

Proof. Sufficiency is clear. To demonstrate necessity, suppose Ψ is calm at $(0, \bar{x}, \bar{z})$ and assume by contradiction the existence of sequences $(x^i, z^i) \to (\bar{x}, \bar{z})$, $(u_1^i, u_2^i) \to (0, 0)$ with $(x^i, z^i) \in \widehat{\Psi}(u_1^i, u_2^i)$ such that

$$d((x^i, z^i), \widehat{\Psi}(0, 0)) > i||(u_1^i, u_2^i)||, \quad \forall i,$$

where $|| \cdot ||$ represents the 1-norm. Put $\tilde{z}^i = z^i + u_1^i$ so that

$$u_2^i \in F(x^i, z^i) - F(x^i, \tilde{z}^i) + F(x^i, \tilde{z}^i) + N_C(\tilde{z}^i).$$

Since F is locally Lipschitz around $(\bar{x}, \bar{z})$, we infer that for i sufficiently large

$$||F(x^i, z^i) - F(x^i, \tilde{z}^i)|| \leq L||u_1^i||,$$

where L is the Lipschitz modulus of F around $(\bar{x}, \bar{z})$. Consequently, for such i, $(x^i, \tilde{z}^i) \in \Psi(u^i)$, with $u^i = u_2^i - F(x^i, z^i) + F(x^i, \tilde{z}^i)$. Clearly, $||u^i|| \leq L||u_1^i|| + ||u_2^i||$ and

$$\begin{aligned} d((x^i, \tilde{z}^i), \Psi(0)) = d((x^i, \tilde{z}^i), \widehat{\Psi}(0, 0)) \geq \\ d((x^i, z^i), \widehat{\Psi}(0, 0)) - ||u_1^i|| > (i-1)||(u_1^i, u_2^i)|| > \frac{(i-1)}{\max\{L, 1\}}||u^i||, \end{aligned}$$

which contradicts the calmness of Ψ at $(0, \bar{x}, \bar{z})$. □

5.2 Calmness of a Class of Perturbation Mappings

In the main result of this section we rely on the results of the following lemma, which is an immediate consequence of Mordukhovich [2006a, Theorem 4.10][2]:

Lemma 5.1. *Let* $Z : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *be a multifunction defined by*

$$Z(x) := \{y \in \mathbb{R}^m | h(x, y) = 0, \ y \in \Omega\},$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ *is a continuously differentiable mapping and* $\Omega \subseteq \mathbb{R}^m$ *is closed. Consider a point* $(\bar{x}, \bar{y}) \in \operatorname{gph} Z$. *If for all* $(x^*, y^*, z^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k$ *the implication*

$$\left.\begin{array}{rcl} x^* & = & \nabla_x^T h(\bar{x}, \bar{y}) z^* \\ y^* & = & -\nabla_y^T h(\bar{x}, \bar{y}) z^* \\ y^* & \in & N_\Omega(\bar{y}) \end{array}\right\} \Longrightarrow x^* = 0 \tag{5.6}$$

[2] Since the well-known Mordukhovich Criterion (see [Rockafellar and Wets, 1998, Chapter 9 Section F.]) states $D^*Z(\bar{x}, \bar{z})(0) = \{0\}$ if and only if Z has the Aubin property at $(\bar{x}, \bar{z}) \in \operatorname{gph} Z$ and it can be verified that (5.6) implies $D^*Z(\bar{x}, \bar{z})(0) = \{0\}$; the result follows.

holds true, then Z has the Aubin property at $(\bar{x}, \bar{y})$.

In the following, we assume, as in the spot market EPEC, that the perturbation mapping Ψ arises from the first order optimality conditions of some perturbed optimization problem. More specifically, we define:

$$\min_z \left\{ f(u_1, z) - u_2^T z \,|\mathcal{A}z + b \in \mathbb{R}^p_- \right\}, \tag{5.7}$$

where $z \in \mathbb{R}^t$, $u_1 \in \mathbb{R}^s$, $u_2 \in \mathbb{R}^t$, $\mathcal{A} \in \mathbb{R}^{p\times t}$, $b \in \mathbb{R}^p$, and assume f is continuously differentiable in both the decision variable z and the parameter u_1. Then the first-order optimality conditions of (5.7) may be written:

$$0 \in \nabla_z f(u_1, z) - u_2 + N_C(z), \tag{5.8}$$

where $C := \{z \in \mathbb{R}^t \,|\mathcal{A}z + b \in \mathbb{R}^p_- \}$. We then define

$$\Psi(u_2) := \left\{ (u_1, z) \in \mathbb{R}^{s+t} \,|u_2 \in \nabla_z f(u_1, z) + N_C(z) \right\}. \tag{5.9}$$

Theorem 5.1. *Let $(\bar{u}_1, \bar{z}, 0)$ be a solution to (5.8) and assume the following conditions hold:*

1. $$\nabla_z f(u_1, z) = \begin{pmatrix} \Delta_1(u_1, z) \\ \Delta_2(z) \end{pmatrix},$$

   where $\Delta_1 \in \mathcal{C}^1(\mathbb{R}^{s+t}; \mathbb{R}^{t_1})$ and $\Delta_2(z) := Dz + c$ with $D \in \mathbb{R}^{t_2\times t}$ such that $t_1 + t_2 = t$ and $c \in \mathbb{R}^{t_2}$.

2. *There exists $\tilde{z}$ such that $\mathcal{A}\tilde{z} + b \in \operatorname{int} \mathbb{R}^p_-$, i.e., there exists a Slater point.*

3. *$\nabla_{u_1} \Delta_1(\bar{u}_1, \bar{z})$ is surjective*

Then the multifunction in (5.9) is calm at $(0, \bar{u}_1, \bar{z})$.

Proof. We begin by defining the following multifunction:

$$\Phi(p_1, p_2) := \left\{ (u_1, z, \lambda) \in \mathbb{R}^{s+t+p} \,\middle|\, \begin{array}{ccc} \Theta_1(u_1, z, \lambda) & = p_1 & \\ \Theta_2(z, \lambda) & = p_2 & \\ \lambda & \in & N_{\mathbb{R}^p_-}(\mathcal{A}z + b) \end{array} \right\},$$

where

$$\begin{aligned} \Theta_1(u_1, z, \lambda) &:= \Delta_1(u_1, z) + \mathcal{A}_1^T \lambda \\ \Theta_2(z, \lambda) &:= \Delta_2(z) + \mathcal{A}_2^T \lambda \end{aligned}$$

such that $\mathcal{A}_1 \in \mathbb{R}^{p\times t_1}$, $\mathcal{A}_2 \in \mathbb{R}^{p\times t_2}$, and $\mathcal{A} = (\mathcal{A}_1 \mid \mathcal{A}_2)$. Since C is polyhedral,

$$N_C(z) = \mathcal{A}^T N_{\mathbb{R}^p_-}(\mathcal{A}z + b).$$

Consequently, by partitioning $u_2 = (u_2^a, u_2^b) \in \mathbb{R}^{t_1} \times \mathbb{R}^{t_2}$, we have

$$\Psi(u_2) = \left\{ (u_1, z) \in \mathbb{R}^{s+t} \,\middle|\, \exists \lambda : (u_1, z, \lambda) \in \Phi(u_2^a, u_2^b) \right\}$$

We now demonstrate, for an arbitrarily fixed $\bar{\lambda}$ with $(0, 0, \bar{u}_1, \bar{z}, \bar{\lambda}) \in \operatorname{gph} \Phi$, that Φ is calm at $(0, 0, \bar{u}_1, \bar{z}, \bar{\lambda})$, after which we show how this implies Ψ is also calm at $(0, \bar{u}_1, \bar{z})$. Start by realizing that $\Phi(p_1, p_2)$ can be written in the following manner:

$$\Phi(p_1, p_2) = \mathcal{S}(p_1) \cap \mathcal{T}(p_2),$$

where

$$\mathcal{S}(p_1) := \{ (u_1, z, \lambda) \,|\, \Theta_1(u_1, z, \lambda) = p_1 \}$$
$$\mathcal{T}(p_2) := \left\{ (u_1, z, \lambda) \,\middle|\, \begin{array}{ccc} \Theta_2(z, \lambda) & = p_2 & \\ \lambda & \in & N_{\mathbb{R}^p_-}(\mathcal{A}z + b) \end{array} \right\}$$

Moreover, note that

$$(\bar{u}_1, \bar{z}, \bar{\lambda}) \in \Phi(0, 0) \Rightarrow (\bar{u}_1, \bar{z}, \bar{\lambda}) \in \mathcal{S}(0) \cap \mathcal{T}(0)$$

Thus, we can show that Φ is calm at $(0, 0, \bar{u}_1, \bar{z}, \bar{\lambda})$ by using the following criteria developed in Klatte and Kummer [2002a, Theorem 3.6]: If the following conditions holds, then Φ is calm at $(0, 0, \bar{u}_1, \bar{z}, \bar{\lambda})$:

$\mathcal{S}$ is calm at $(0, \bar{u}_1, \bar{z}, \bar{\lambda})$
$\mathcal{T}$ is calm at $(0, \bar{u}_1, \bar{z}, \bar{\lambda})$
$\mathcal{S}^{-1}$ has the Aubin property at $(\bar{u}_1, \bar{z}, \bar{\lambda}, 0)$
$\mathcal{S} \cap \mathcal{T}(0)$ is calm at $(0, \bar{u}_1, \bar{z}, \bar{\lambda})$

Since the multifunction $\mathcal{T}$ is polyhedral, it clearly follows from Robinson [1981, Proposition 1] that it is calm[3] at $(0, \bar{u}_1, \bar{z}, \bar{\lambda})$. Thus, we show the remaining three conditions hold, from which the desired result follows

1. $\mathcal{S}$ has the Aubin property at $(0, \bar{u}_1, \bar{z}, \bar{\lambda})$.

2. $\mathcal{S} \cap \mathcal{T}(0)$ has the Aubin property at $(0, \bar{u}_1, \bar{z}, \bar{\lambda})$.

3. $\mathcal{S}^{-1}$ has the Aubin property at $(\bar{u}_1, \bar{z}, \bar{\lambda}, 0)$.

Recall Lemma 5.1. For 1. and 2. put $x := p_1$,$y := (u_1, z, \lambda)$, $h(x, y) := \Theta_1(y) - x$, and $\Omega := \mathbb{R}^m$ for 1. and $\Omega := \mathcal{T}(0)$ for 2., respectively. Then $Z = \mathcal{S}$ in 1. and $Z = \mathcal{S} \cap \mathcal{T}(0)$ in 2. The first relation in the assumption of (5.6) then yields $x^* = -z^*$. Hence, the second relation amounts to $y^* = \nabla_y^T h(\bar{x}, \bar{y}) x^*$. By using the partition $y^* = (u_1^*, z^*, \lambda^*)$, the first component of the previous relation, in

[3]Actually, in Robinson [1981], polyhedral multifunctions are shown to be *upper Lipschitzian*, a stronger property, which we make use of later in this proof

accordance with that of y, now reads

$$u_1^* = \nabla_{u_1}^T h(\bar{x}, \bar{y})x^* = \nabla_{u_1}^T \Theta_1(\bar{u}_1, \bar{z}, \bar{\lambda})x^*.$$

By observing that $\mathcal{T}(0) = \mathbb{R}^s \times \mathcal{T}'$ for some subset $\mathcal{T}' \subseteq \mathbb{R}^t \times \mathbb{R}^p$, we have in both cases 1. and 2. that $\Omega = \mathbb{R}^s \times \mathcal{T}'$. Consequently, the third relation in the assumption (5.6) yields $u_1^* = 0$ in either case and hence, upon using the explicit structure of Θ_1, (5.6) reduces to

$$\nabla_{u_1}^T \Delta_1(\bar{u}_1, \bar{z})x^* = 0 \Rightarrow x^* = 0.$$

This of course follows immediately from assumption 3. of the current theorem. For the proof of 3., put $y := p_1$, $x = (u_1, z, \lambda)$, $h(x, y) := \Theta_1(x) - y$, and $\Omega := \mathbb{R}$. Then $Z = S^{-1}$ and the assumption of (5.6) trivially reduces to the implication $x^* = 0$.

Summarizing, we have shown that Φ is calm at $(0, 0, \bar{u}_1, \bar{z}, \bar{\lambda})$. That is, there exists a constant $L_{\bar{\lambda}} > 0$ and neighborhoods $\mathcal{U}$ of $(\bar{u}_1, \bar{z})$, $\mathcal{V}$ of $\bar{\lambda}$, and $\mathcal{W}$ of $(0, 0)$ such that:

$$d((u_1, z, \lambda), \Phi(0, 0)) \leq L_{\bar{\lambda}}||(p_1, p_2)||, \quad \begin{array}{c} \forall (u_1, z, \lambda) \in [\mathcal{U} \times \mathcal{V}] \cap \Phi(p_1, p_2) \\ \forall (p_1, p_2) \in \mathcal{W} \end{array}$$

In particular, there exists an open ball $\mathbb{B}^\circ_{\varepsilon_{\bar{\lambda}}}(\bar{\lambda}) > 0$ with radius $\varepsilon_{\bar{\lambda}} > 0$ such that:

$$\begin{aligned} \forall (u_1, z, \lambda) \in [\mathbb{B}^\circ_{\varepsilon_{\bar{\lambda}}}(\bar{u}_1, \bar{z}) \times \mathbb{B}^\circ_{\varepsilon_{\bar{\lambda}}}(\bar{\lambda})] \cap \Phi(p_1, p_2), \forall (p_1, p_2) \in \mathbb{B}^\circ_{\varepsilon_{\bar{\lambda}}}(0, 0) : \\ d((u_1, z, \lambda), \Phi(0, 0)) \leq L_{\bar{\lambda}}||(p_1, p_2)||. \end{aligned} \tag{5.10}$$

As $\bar{\lambda}$ was chosen arbitrarily, the previous argument indicates that for all Lagrange multipliers λ with $(\bar{u}_1, \bar{z}, \lambda) \in \Phi(0, 0)$ we obtain constants L_λ, ε_λ such that (5.10) holds at $(\bar{u}_1, \bar{z}, \lambda)$ by replacing $L_{\bar{\lambda}}$ with L_λ and $\varepsilon_{\bar{\lambda}}$ with ε_λ. Consider now the multifunction that assigns to each (u_1, z, u_2) the set of Lagrange multipliers associated with the optimization problem (5.7):

$$\Lambda(u_1, z, u_2) := \left\{ \lambda \in \mathbb{R}^p \;\middle|\; \begin{array}{c} \nabla_z f(u_1, z) - u_2 + \mathcal{A}^T \lambda = 0 \\ \mathcal{A}z + b \in \mathbb{R}^p_- \\ \lambda \in N_{\mathbb{R}^p_-}(\mathcal{A}z + b) \end{array} \right\}$$

and note that

$$\lambda \in \Lambda(u_1, z, u_2) \Leftrightarrow (u_1, z, \lambda) \in \Phi(u_2^a, u_2^b).$$

Due to the fact that in the previous argument $\bar{\lambda}$ was arbitrarily chosen, we may infer that for all $\lambda \in \Lambda(\bar{u}_1, \bar{z}, 0)$, there exist constants $\varepsilon_\lambda > 0$ and $L_\lambda > 0$ such that (5.10) holds.

By taking the open balls $\mathbb{B}^\circ_{\varepsilon_\lambda}(\lambda)$ introduced above with $\lambda \in \Lambda(\bar{u}_1, \bar{z}, 0)$, we can provide an open covering of $\Lambda(\bar{u}_1, \bar{z}, 0)$:

$$\bigcup_{\lambda \in \Lambda(\bar{u}_1, \bar{z}, 0)} \mathbb{B}^\circ_{\varepsilon_\lambda}(\lambda) \supset \Lambda(\bar{u}_1, \bar{z}, 0),$$

Given the Slater point assumption and the fact that C is convex, we know $\Lambda(\bar{u}_1, \bar{z}, 0)$ is a nonempty compact set (see e.g., Bonnans and Shapiro [2000, Theorem 3.6]). Thus, the open covering contains a finite open subcovering, which in particular means that there exist $\lambda_i \in \Lambda(\bar{u}_1, \bar{z}, 0)$ for $i = 1, \ldots, \kappa$, such that

$$\bigcup_{i=1}^{\kappa} \mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(\lambda_i) = \widehat{\mathcal{O}} \supset \Lambda(\bar{u}_1, \bar{z}, 0).$$

Accordingly, we have for all $i = 1, \ldots, \kappa$

$$\begin{aligned} &d((u_1, z, \lambda), \Phi(0,0)) \leq L_{\lambda_i}||(p_1, p_2)|| \\ &\quad \forall (u_1, z, \lambda) \in [\mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(\bar{u}_1, \bar{z}) \times \mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(\lambda_i)] \cap \Phi(p_1, p_2) \quad \forall (p_1, p_2) \in \mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(0,0) \end{aligned} \tag{5.11}$$

By letting

$$\varepsilon = \min_{1,\ldots,\kappa} \varepsilon_{\lambda_i} \quad L = \max_{1,\ldots,\kappa} L_{\lambda_i},$$

we can demonstrate that Φ satisfies the following uniform calmness property

$$\begin{aligned} &d((u_1, z, \lambda), \Phi(0,0)) \leq L||(p_1, p_2)|| \\ &\qquad \forall (u_1, z, \lambda) \in [\mathbb{B}^\circ_{\varepsilon}(\bar{u}_1, \bar{z}) \times \widehat{\mathcal{O}}] \cap \Phi(p_1, p_2) \quad \forall (p_1, p_2) \in \mathbb{B}^\circ_{\varepsilon}(0,0) \end{aligned} \tag{5.12}$$

Indeed, by choosing an arbitrary $(u_1, z, \lambda, p_1, p_2)$ from the neighborhoods indicated in (5.12), we see that by the definition of $\widehat{\mathcal{O}}$, there exists $i \in \{1, \ldots, \kappa\}$ such that $\lambda \in \mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(\lambda_i)$. Moreover, by the definition of ε, we have $(p_1, p_2) \in \mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(0,0)$ and $(u_1, z) \in \mathbb{B}^\circ_{\varepsilon_{\lambda_i}}(\bar{u}_1, \bar{z})$. By (5.11) and the definition of L, we have that

$$d((u_1, z, \lambda), \Phi(0,0)) \leq L_{\lambda_i}||(p_1, p_2)|| \leq L||(p_1, p_2)||$$

Whence, (5.12).

Now that we have established a uniform calmness condition for Φ, we can show that Ψ is calm at $(0, \bar{u}_1, \bar{z})$.

We know that the Slater point assumption along with the convexity of C is equivalent to the (MFCQ) holding at $\bar{z} \in C$ (see Bonnans and Shapiro [2000, Corollary 2.101 and Proposition 2.104]). As a consequence of the polyhedrality of $N_{\mathbb{R}^p_-}(\cdot)$, we may invoke Bonnans and Shapiro [2000, Lemma 4.44] via Bonnans and Shapiro [2000, remark 4.45] in order to show that Λ is upper-Lipschitz at $(\bar{u}_1, \bar{z}, 0)$, which by definition implies in particular that there exists $\hat{\varepsilon} > 0$ such that for all $(u_1, z, u_2) \in \mathbb{B}^\circ_{\hat{\varepsilon}}(\bar{u}_1, \bar{z}) \times \mathbb{B}^\circ_{\hat{\varepsilon}}(0)$,

$$d(\lambda, \Lambda(\bar{u}_1, \bar{z}, 0)) \leq \hat{L}||(u_1, z, u_2) - (\bar{u}_1, \bar{z}, 0)|| \ \forall \lambda \in \Lambda(u_1, z, u_2).$$

Moreover, the compactness of $\Lambda(\bar{u}_1, \bar{z}, 0)$ implies the existence of an $\varepsilon' > 0$ such that

$$\Lambda(\bar{u}_1, \bar{z}, 0) + B^\circ_{\varepsilon'}(0) \subseteq \widehat{\mathcal{O}}.$$

By letting $\tilde{\varepsilon} = \min\{\hat{\varepsilon}, \varepsilon'/\hat{L}, \varepsilon\}$, we can argue for an arbitrary $u_2 \in \mathbb{B}^\circ_{\tilde{\varepsilon}}(0,0)$ that the following holds

$$(u_1, z) \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(\bar{u}_1, \bar{z}) \cap \Psi(u_2) \Rightarrow \exists \lambda \in \widehat{\mathcal{O}} : (u_1, z, \lambda) \in \Phi(u_2^a, u_2^b) \tag{5.13}$$

Indeed, given $(u_1, z) \in \Psi(u_2)$, by definition there must exist λ such that $(u_1, z, \lambda) \in \Phi(u_2^a, u_2^b)$, which, as noted above, implies that $\lambda \in \Lambda(u_1, z, u_2)$. Then since $(u_1, z) \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(\bar{u}_1, \bar{z})$ and $u_2 \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(0)$, $d(\lambda, \Lambda(\bar{z}, \bar{u}_1, 0)) \leq \varepsilon'$. Thus, $\lambda \in \Lambda(\bar{u}_1, \bar{z}, 0) + \mathbb{B}^{\circ}_{\varepsilon'}(0) \subseteq \widehat{\mathcal{O}}$.

Then by taking an arbitrary $(u_1, z) \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(\bar{u}_1, \bar{z}) \cap \Psi(u_2)$ and $u_2 \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(0, 0)$, (5.13) implies that there exists a $\lambda \in \widehat{\mathcal{O}}$ such that $(u_1, z, \lambda) \in \Phi(u_2^a, u_2^b)$. Suppose now that $(\tilde{u}_1, \tilde{z}, \tilde{\lambda}) \in \Phi(0, 0)$ is the point such that $d((u_1, z, \lambda), \Phi(0, 0)) = ||(u_1, z, \lambda) - (\tilde{u}_1, \tilde{z}, \tilde{\lambda})||$. Clearly, $(\tilde{u}_1, \tilde{z}) \in \Psi(0)$. Then by (5.12):

$$\begin{aligned} d((u_1, z), \Psi(0)) \leq ||(u_1, z) - (\tilde{u}_1, \tilde{z})|| \leq ||(u_1, z, \lambda) - (\tilde{u}_1, \tilde{z}, \tilde{\lambda})|| = \\ d((u_1, z, \lambda), \Phi(0, 0)) \leq L||(u_2^a, u_2^b)||. \end{aligned} \tag{5.14}$$

Thus, we have the following

$$d((u_1, z), \Psi(0)) \leq L||(u_2^a, u_2^b)|| \quad \forall (u_1, z) \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(\bar{u}_1, \bar{z}) \cap \Psi(u_2), \quad \forall u_2 \in \mathbb{B}^{\circ}_{\bar{\varepsilon}}(0).$$

Whence we know Ψ is calm at $(0, \bar{u}_1, \bar{z})$. $\square$

Chapter 6

Coderivative Transformation Formulae for Normal Cone Mappings

As seen in the previous chapters, the ability to explicitly calculate the coderivative to the normal cone mapping to a system of equalities and inequalities plays a crucial role in our analysis. In this chapter, we provide various results on coderivative transformation formulae, some of which first appeared in Mordukhovich and Outrata [2001, 2007], Henrion and Römisch [2007] and Henrion et al. [2009c]. The important results from the latter paper, which the author c-owrote, are provided with proofs, thus providing a more complete presentation.

6.1 Polyhedral Feasible Sets

Throughout this section, we will consider the normal cone mapping $N_{\mathcal{C}}$, where $\mathcal{C}$ is a general polyhedron defined

$$\mathcal{C} := \left\{ z \in \mathbb{R}^t \,|\mathcal{A}z \leq b \right\}.$$

Here, $\mathcal{A}$ is a (p,t)-matrix, $b \in \mathbb{R}^p$ and the inequality is understood componentwise. We will continue to use $I(\bar{z})$, $I_+(\bar{z},\bar{\lambda})$, $I_0(\bar{z},\bar{\lambda})$, and $L(\bar{z})$ to denote the set of active, strongly active, weakly active, and inactive indices, leaving off the arguments when it is clear in context. Initial results on transformation formulae for the cases in which $\mathcal{C}$ is an orthant or rectangle were provided in Dontchev and Rockafellar [1996] and Outrata [2001]. As it plays an important role throughout this text, we provide the following result (see Henrion and Römisch [2007, Corollary 3.5] or Henrion et al. [2009c, Corollary 3.1]); the term 'regular' refers to the surjectivity of $\mathcal{A}$.

Proposition 6.1 (regular systems of linear inequalities). *For $\mathcal{C}$ as defined above, let $v^* \in \mathbb{R}^t$, $(\bar{z},\bar{v}) \in \operatorname{gph} N_{\mathcal{C}}$, and assume $\operatorname{rank}\mathcal{A} = p$ and $\mathcal{A}\bar{z} = b$. Then*

$$D^*N_{\mathcal{C}}(\bar{z},\bar{v})(v^*) = \begin{cases} \mathcal{A}^T w^*, & \mathcal{A}_j v^* = 0,\ \forall j : \bar{\lambda}_j > 0 \\ \emptyset & \textit{otherwise} \end{cases}$$

where

$$w_j^* = 0,\ \forall j : \bar{\lambda}_j = 0, \mathcal{A}_j v^* < 0$$
$$w_j^* \geq 0,\ \forall j : \bar{\lambda}_j = 0, \mathcal{A}_j v^* > 0$$

and $\bar{\lambda} \in \mathbb{R}_+^p$ *is the unique Lagrange multiplier defined via* $\bar{v} = \mathcal{A}^T \bar{\lambda}$*.*

We now provide a result useful in situations where equalities are involved and one in which transforming the equalities into inequalities would destroy certain regularity properties. We define the set

$$\widetilde{C} = \left\{ z \in \mathbb{R}^t \,\middle|\, \mathcal{A}^1 z \leq 0, \mathcal{A}^2 z = 0 \right\},$$

where $\mathcal{A}^1 \in \mathbb{R}^{p_- \times t}$ and $\mathcal{A} \in \mathbb{R}^{p_0 \times t}$ such that $p = p_- + p_0$ and $\mathcal{A} = \begin{pmatrix} \mathcal{A}^1 \\ \mathcal{A}^2 \end{pmatrix}$.

Corollary 6.1 (regular systems of linear equalities and inequalities). *For* $\widetilde{C}$ *as defined above, let* $v^* \in \mathbb{R}^t$, $(\bar{z}, \bar{v}) \in \operatorname{gph} N_{\widetilde{C}}$*, and assume* $\operatorname{rank} \mathcal{A} = p$ *and that* $\mathcal{A}^1 \bar{z} = 0$*. Then*

$$D^* N_{\widetilde{C}}(\bar{z}, \bar{v})(v^*) = \begin{cases} \mathcal{A}^T w^*, & \begin{array}{l} \mathcal{A}_j^1 v^* = 0,\ \forall j \in \{1, \ldots, p_-\} : \bar{\lambda}_j > 0 \\ \mathcal{A}^2 v^* = 0 \end{array} \\ \emptyset & \textit{otherwise} \end{cases}$$

where

$$w_j^* = 0,\ \forall j \in \{1, \ldots, p_-\} : \bar{\lambda}_j = 0, \mathcal{A}_j v^* < 0$$
$$w_j^* \geq 0,\ \forall j \in \{1, \ldots, p_-\} : \bar{\lambda}_j = 0, \mathcal{A}_j v^* > 0$$

and $\bar{\lambda}$ *is the unique Lagrange multiplier defined via* $\bar{v} = \mathcal{A}^T \bar{\lambda}$ *such that* $\bar{\lambda}_j \in \mathbb{R}_+$ *for all* $j = 1, \ldots, p_-$ *and* $\bar{\lambda}_j \in \mathbb{R}$ *for all* $j = p_- + 1, \ldots, p_0$*.*

Proof. First note that we may write the normal cone at $\bar{z}$ in the following way

$$N_{\widetilde{C}}(\bar{z}) = \left\{ \bar{v} = \mathcal{A}^T \bar{\lambda} \,\middle|\, \bar{\lambda} \in N_{\mathbb{R}_-^{p_-} \times \{0\}}(\mathcal{A}\bar{z}) \right\}.$$

Here, $\{0\}$ represents the zero vector in $\mathbb{R}^{p_0}$. Then from Mordukhovich and Outrata [2001, Theorem 3.4] or equivalently Mordukhovich [2006a, Theorem 1.127], we have that

$$D^* N_{\widetilde{C}}(\bar{z}, \bar{v})(v^*) = \mathcal{A}^T D^* N_{\mathbb{R}_-^{p_-} \times \{0\}}(\mathcal{A}\bar{z}, \bar{\lambda})(\mathcal{A} v^*). \tag{6.1}$$

Partition $\bar{\lambda} = (\bar{\lambda}^1, \bar{\lambda}^2) \in \mathbb{R}^{p_-} \times \mathbb{R}^{p_0}$, then by definition

$$D^* N_{\mathbb{R}^{p_-}_- \times \{0\}}(\mathcal{A}\bar{z}, \bar{\lambda})(\mathcal{A}v^*) = \Big\{(z_1^*, z_2^*) \in \mathbb{R}^{p_-} \times \mathbb{R}^{p_0} \mid \\ (z_1^*, z_2^*, -\mathcal{A}^1 v^*, -\mathcal{A}^2 v^*) \in N_{\operatorname{gph} N_{\mathbb{R}^{p_-}_- \times \{0\}}}(\mathcal{A}^1\bar{z}, \mathcal{A}^2\bar{z}, \bar{\lambda}^1, \bar{\lambda}^2)\Big\}.$$

Next, using elementary properties of the normal cone mapping (cf. Rockafellar and Wets [1998, Proposition 6.41]), it easy is to show that

$$(h^1, h^2, w^1, w^2) \in \operatorname{gph} N_{\mathbb{R}^{p_-}_- \times \{0\}} \Leftrightarrow (h^1, w^1, h^2, w^2) \in \operatorname{gph} N_{\mathbb{R}^{p_-}_-} \times [\{0\} \times \mathbb{R}^{p_0}] .$$

Utilizing this last fact, we may write

$$D^* N_{\mathbb{R}^{p_-}_- \times \{0\}}(\mathcal{A}\bar{z}, \bar{\lambda})(\mathcal{A}v^*) = \Big\{(z_1^*, z_2^*) \mid \\ (z_1^*, -\mathcal{A}^1 v^*, z_2^*, -\mathcal{A}^2 v^*) \in N_{\operatorname{gph} N_{\mathbb{R}^{p_-}_-}}(\mathcal{A}^1\bar{z}, \bar{\lambda}^1) \times N_{[\{0\} \times \mathbb{R}^{p_0}]}(\mathcal{A}^2\bar{z}, \bar{\lambda}^2)\Big\},$$

which leads us to the following conclusion

$$D^* N_{\mathbb{R}^{p_-}_- \times \{0\}}(\mathcal{A}\bar{z}, \bar{\lambda})(\mathcal{A}v^*) = \\ D^* N_{\mathbb{R}^{p_-}_-}(\mathcal{A}^1\bar{z}, \bar{\lambda}^1)(\mathcal{A}^1 v^*) \times D^* N_{\{0\}}(\mathcal{A}^2\bar{z}, \bar{\lambda}^2)(\mathcal{A}^2 v^*). \quad (6.2)$$

Thus it remains for us to calculate the two simpler coderivatives. Note first that since $N_{[\{0\} \times \mathbb{R}^{p_0}]}(\mathcal{A}^2\bar{z}, \bar{\lambda}^2) = \mathbb{R}^{p_0} \times \{0\}$,

$$D^* N_{\{0\}}(\mathcal{A}^2\bar{z}, \bar{\lambda}^2)(\mathcal{A}^2 v^*) = \begin{cases} \mathbb{R}^{p_0}, & \text{if } \mathcal{A}^2 v^* = 0 \\ \emptyset & \text{otherwise} \end{cases} . \quad (6.3)$$

Moreover, Proposition 6.1 states

$$D^* N_{\mathbb{R}^{p_-}_-}(\mathcal{A}^1\bar{z}, \bar{\lambda}^1)(\mathcal{A}^1 v^*) = \begin{cases} w^* & \mathcal{A}^1_j v^* = 0, \ \forall j : \bar{\lambda}_j > 0 \\ \emptyset & \text{otherwise} \end{cases} , \quad (6.4)$$

with

$$w_j^* = 0, \ \forall j \in \{1, \ldots, p_-\} : \bar{\lambda}_j = 0, \mathcal{A}_j v^* < 0 \\ w_j^* \geq 0, \ \forall j \in \{1, \ldots, p_-\} : \bar{\lambda}_j = 0, \mathcal{A}_j v^* > 0.$$

Then substituting (6.3) and (6.4) into (6.2) and referring back to the initial expression (6.1) we obtain the desired result. □

Example 6.1 (calculating $D^* N_{K(\bar{x},\bar{z})}(0, 0)(v^*)$ with LICQ). *For a solution $(\bar{x}, \bar{z})$ to the generalized equation (4.2), assume $A(\bar{z}) = 0$ and $\nabla A(\bar{z})$ is surjective.*

Here, the critical cone to C at $(\bar{z}, F(\bar{x}, \bar{z}))$ becomes

$$K(\bar{x}, \bar{z}) = \left\{h \,\middle|\, \nabla A_j(\bar{z})h \leq 0 \, (j \in I_0(\bar{z}, \bar{\lambda})), \nabla A_j(\bar{z})h = 0, \, (j \in I_+(\bar{z}, \bar{\lambda}))\right\},$$

where $\bar{\lambda}$ is the uniquely defined Lagrange multiplier associated with the equation: $F(\bar{x}, \bar{z}) = -\nabla^T A(\bar{z})\bar{\lambda}$. Then Corollary 6.1 states

$$D^* N_{K(\bar{x},\bar{z})}(0,0)(v^*) = \begin{cases} \nabla^T A(\bar{z})w^*, & \begin{array}{l} \bar{\mu}_j \nabla A_j(\bar{z})v^* = 0, j \in I_0(\bar{z}, \bar{\lambda}) \\ \nabla A_j(\bar{z})v^* = 0, \; j \in I_+(\bar{z}, \bar{\lambda}) \end{array} \\ \emptyset & \text{otherwise}, \end{cases}$$

with

$$\begin{aligned} w_j^* &= 0, \; \forall j : \bar{\mu}_j = 0, \bar{\lambda}_j = 0, \nabla A_j(\bar{z})v^* < 0 \\ w_j^* &\geq 0, \; \forall j : \bar{\mu}_j = 0, \bar{\lambda}_j = 0, \nabla A_j(\bar{z})v^* > 0 \end{aligned}$$

and $\bar{\mu}$ defined via $0 = \nabla^T A_{I_0}(\bar{z})\bar{\mu}$. However, $\nabla A_{I_0}(\bar{z})$ is surjective, as it is a submatrix of a surjective matrix, hence $\bar{\mu} = 0$, in which case

$$D^* N_{K(\bar{x},\bar{z})}(0,0)(v^*) = \begin{cases} \nabla^T A(\bar{z})w^*, & \nabla A_j(\bar{z})v^* = 0, \; j \in I_+(\bar{z}, \bar{\lambda}) \\ \emptyset & \text{otherwise}, \end{cases} \tag{6.5}$$

and

$$\begin{aligned} w_j^* &= 0, \; \forall j : \bar{\lambda}_j = 0, \nabla A_j(\bar{z})v^* < 0 \\ w_j^* &\geq 0, \; \forall j : \bar{\lambda}_j = 0, \nabla A_j(\bar{z})v^* > 0 \end{aligned}$$

Referring back to Proposition 6.1, it is easy to see that

$$D^* N_{K(\bar{x},\bar{z})}(0,0)(v^*) = \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z})v^*).$$

In the event $\mathcal{A}$ is not surjective, it is still possible to obtain an exact formula for $D^* N_{\mathcal{C}}$ More precisely, we have the following theorem (Henrion and Römisch [2007, Proposition 3.2] and Henrion et al. [2009c, Theorem 3.2])

Theorem 6.1 (**nonregular polyhedra**). *Define $\mathcal{C} := \{z \in \mathbb{R}^t \,|\, \mathcal{A}z \leq b\}$, where $b \in \mathbb{R}^p$ and $\mathcal{A}$ is a matrix of order (p, t). Let $(\bar{z}, \bar{v}) \in \operatorname{gph} N_{\mathcal{C}}$ and assume without loss of generality that $\mathcal{A}\bar{z} = b$. Let $\bar{\lambda} \in \mathbb{R}^p_+$ be defined by the relation $\mathcal{A}^T\bar{\lambda} = \bar{v}$. Then,*

$$D^* N_{\mathcal{C}}(\bar{z}, \bar{v})(v^*) = \left\{x^* \,\middle|\, (x^*, -v^*) \in \bigcup_{I_+(\bar{z},\bar{\lambda}) \subseteq I_1 \subseteq I_2 \subseteq \{1,\ldots,p\}} P_{I_1,I_2} \times Q_{I_1,I_2}\right\}, \tag{6.6}$$

where

$$P_{I_1,I_2} = \text{con}\left\{\mathcal{A}_j^T | j \in \chi\left(I_2\right) \backslash I_1\right\} + \text{span}\left\{\mathcal{A}_j^T | j \in I_1\right\}$$
$$Q_{I_1,I_2} = \left\{h \in \mathbb{R}^t | \mathcal{A}_j h = 0 \ \ (j \in I_1)\,,\ \mathcal{A}_j h \le 0 \quad (j \in \chi\left(I_2\right) \backslash I_1)\right\}$$

and for $(I' \subseteq \{1, \ldots, p\})$

$$\chi(I') := \{j \in \{1, \ldots, p\} \mid \\ \textit{if } \mathcal{A}_j h \le 0 \textit{ for } j \in \{1, \ldots, p\} \setminus I' \textit{and } \mathcal{A}_j h = 0 \textit{ for } j \in I',\ \textit{then } \mathcal{A}_j h = 0\}$$

Note that the existence of $\bar{\lambda}$ in Theorem 6.1 is guaranteed by the polyhedrality of $\mathcal{C}$. Though using (6.6) may a somewhat cumbersome task, it indicates that it is possible to obtain an exact formula even in situations where $\mathcal{A}$ is not surjective. In addition, it is possible to obtain an upper approximation of $D^*N_{\mathcal{C}}$ (see Henrion and Römisch [2007, Corollary 3.4] and Henrion et al. [2009c, Corollary 3.3]), the accuracy of which is discussed at length in Henrion et al. [2009c, Section 3]. Lastly, we mention that there exists a generalization of Theorem 6.1 in infinite dimensions for the case of finitely many linear inequalities[1] in reflexive Banach spaces (see Henrion et al. [2009a, Theorem 4.1]).

6.2 Nonlinearly Constrained Sets

Starting with the regular case, we now move on to the nonlinear setting and begin by providing the following result (see Mordukhovich and Outrata [2001, Theorem 3.4/Remark 3.5], Mordukhovich [2006a, Theorem 1.127], or Henrion et al. [2009c, Theorem 3.1]).

Theorem 6.2 (**a transformation formula for** D^*N_C **using LICQ**)**.** *Let* $C = A^{-1}(P)$*, where* $A : \mathbb{R}^t \to \mathbb{R}^p$ *is twice continuously differentiable and* $P \subseteq \mathbb{R}^p$ *is some closed subset. Consider points* $\bar{z} \in C$ *and* $\bar{v} \in N_C(\bar{z})$*. If the Jacobian* $\nabla A(\bar{z})$ *is surjective, then*

$$D^*N_C(\bar{z}, \bar{v})(v^*) = \\ \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z})\right) v^* + \nabla^T A\left(\bar{z}\right) D^*N_P\left(A(\bar{z}), \bar{\lambda}\right)\left(\nabla A\left(\bar{z}\right) v^*\right) \quad (6.7)$$

Here, the A_j *are the components of* A *and* $\bar{\lambda}$ *is the unique solution of the equation* $\nabla^T A(\bar{z})\bar{\lambda} = \bar{v}$*, i.e.,*

$$\bar{\lambda} = \left(\nabla A(\bar{z}) \nabla^T A(\bar{z})\right)^{-1} \nabla A(\bar{z}) \bar{v}.$$

Theorem 6.2 effectively allows us to transfer the difficulty of calculating D^*N_C onto D^*N_P, given P is some closed polyhedron, which we now know to be a much

[1] That is, inequalities of the form: $\langle x^*, x\rangle \le 0$ with $x \in X$, $x^* \in X^*$, where X and X^* represent a Banach space and its dual.

simpler calculation. Moreover, in the case where $P = \mathbb{R}^p_-$, one can easily develop a transformation formula in which the righthand side contains no general terms using Proposition 6.1 or Corollary 6.1. Thus, one can again obtain an explicit characterization of dom D^*N_C, even in a nonlinear case.

The following theorem first appeared in Mordukhovich and Outrata [2007, Theorem 4.1] and is a generalization of Theorem 6.2. As it is an upper-approximation, it is difficult to say how much bigger the right-hand side is than the actual coderivative itself when surjectivity of $\nabla A(\bar{z})$ is not present. Nevetheless, the constraint qualifications involved are much weaker than surjectivity and thus, there is a better chance they hold, thereby providing the user with an approximation of the coderivative. For a more in depth discussion of this topic, the reader is directed to Henrion et al. [2009c, Section 3 Examples 3.1 & 3.2].

Theorem 6.3 (a transformation formula for D^*N_C using MFCQ and Calmness). *Consider the set $C = \{z \in \mathbb{R}^t \mid A_j(z) \le 0 \ (j = 1, \dots, p)\}$, where $A : \mathbb{R}^t \to \mathbb{R}^p$ is twice continuously differentiable. Fix some $\bar{z} \in C$ and $\bar{v} \in N_C(\bar{z})$ such that, without loss of generality, $A(\bar{z}) = 0$ and suppose that the following two constraint qualifications are fulfilled:*

1. *MFCQ holds at $\bar{z}$*

2. *The multifunction*

$$M(\vartheta) := \{(z, \lambda) \,|\, (A(z), \lambda) + \vartheta \in \operatorname{gph} N_{\mathbb{R}^p_-}\}$$

   is calm at $\left(0, \bar{z}, \bar{\lambda}\right)$ for all $\bar{\lambda} \ge 0$ with $\nabla^T A(\bar{z})\bar{\lambda} = \bar{v}$.

Then,

$$D^*N_C(\bar{z}, \bar{v})(v^*) \subseteq \bigcup_{\substack{\bar{\lambda} \ge 0 \\ \nabla^T A(\bar{z})\bar{\lambda} = \bar{v}}} \left\{ \left(\sum_{j=1}^{p} \bar{\lambda}_i \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z})\, D^*N_{\mathbb{R}^p_-} \left(0, \bar{\lambda}\right) \left(\nabla A(\bar{z})\, v^*\right) \right\}.$$

Assumption 2. in Theorem 6.3 can be quite difficult to verify in general. Therefore, we develop a replacement constraint qualification using only primal variables. The following result demonstrates that the fulfillment of a certain primal condition implies the calmness of the multifunction M at $(0, \bar{z}, \bar{\lambda})$.

Proposition 6.2 (a primal calmness condition). *If for all nonempty subsets $I \subseteq \{1, \dots, m\}$ the multifunctions*

$$H_I(\alpha) = \{z \mid A_i(z) = \alpha_i \ (i \in I), \ A_i(z) \le 0 \ (i \in I^c)\}$$

are calm at $(0, \bar{z})$, then the multifunction M introduced in Theorem 6.3 is calm at $\left(0, \bar{z}, \bar{\lambda}\right)$ for any $\bar{\lambda}$ specified there.

Proof. Throughout this proof we use the 1−norm of vectors. Note first, that for $I = \emptyset$, H_I is trivially calm as a constant multifunction. Hence, this special case can be excluded from the assumption. Next, observe that, by $A(\bar{z}) = 0$, one has indeed $(0, \bar{z}) \in \operatorname{gph} H_I$ for all $I \subseteq \{1, \ldots, p\}$. The calmness assumption means that for any $I \subseteq \{1, \ldots, p\}$, there exist constants $\delta_I, \varepsilon_I, L_I > 0$ such that

$$d(z, H_I(0)) \leq L_I \|\alpha\| \quad \forall z \in \mathbb{B}_{\delta_I}(\bar{z}) \cap H_I(\alpha)\ \forall \alpha : \alpha_i \in (-\varepsilon_I, \varepsilon_I)\ \ (i \in I).$$

Putting

$$\delta := \min_{I \subseteq \{1,\ldots,p\}} \delta_I, \quad \varepsilon := \min_{I \subseteq \{1,\ldots,p\}} \varepsilon_I, \quad L := \max_{I \subseteq \{1,\ldots,p\}} L_I,$$

one obtains that $\delta, \varepsilon, L > 0$ and

$$d(z, H_I(0)) \leq L \|\alpha\|, \quad \begin{array}{l} \forall z \in \mathbb{B}_\delta(\bar{z}) \cap H_I(\alpha) \\ \forall \alpha : \alpha_i \in (-\varepsilon, \varepsilon)\ \ (i \in I) \\ \forall I \subseteq \{1, \ldots, p\}. \end{array} \tag{6.8}$$

Due to $A(\bar{z}) = 0$, we may further shrink $\delta > 0$ such that

$$|A_j(z)| \leq \varepsilon, \quad \forall z \in \mathbb{B}_\delta(\bar{z})\ \forall j \in \{1, \ldots, p\}. \tag{6.9}$$

Now, consider any $\bar{\lambda} \geq 0$. Then, $\bar{\lambda} \in N_{\mathbb{R}^p_-}(0)$ and so $\left(\bar{z}, \bar{\lambda}\right) \in M(0)$. We show that

$$d((z, \lambda), M(0)) \leq (L + 1) \|\vartheta\|, \quad \begin{array}{l} \forall (z, \lambda) \in M(\vartheta) \cap (\mathbb{B}_\delta(\bar{z}) \times \mathbb{R}^p) \\ \forall \vartheta = (\vartheta_1, \vartheta_2) \in \mathbb{B}_\varepsilon(0) \times \mathbb{R}^p \end{array}. \tag{6.10}$$

This would prove the asserted calmness of M at $\left(0, \bar{z}, \bar{\lambda}\right)$. To this aim, choose arbitrary $\vartheta = (\vartheta_1, \vartheta_2) \in \mathbb{B}_\varepsilon(0) \times \mathbb{R}^p$ and $(z, \lambda) \in M(\vartheta) \cap (\mathbb{B}_\varepsilon(\bar{z}) \times \mathbb{R}^p)$. Note first that $(z, \lambda) \in M(\vartheta)$ amounts to $\lambda + \vartheta_2 \in N_{\mathbb{R}^p_-}(A(z) + \vartheta_1)$. Accordingly,

$$A(z) + \vartheta_1 \leq 0,\ \lambda + \vartheta_2 \geq 0,\ (\lambda_j + \vartheta_{2j})(A_j(z) + \vartheta_{1j}) = 0\ \forall j \in \{1, \ldots, p\}. \tag{6.11}$$

For the fixed z, define

$$I_z := \{j \in \{1, \ldots, p\} | A_j(z) + \vartheta_{1j} = 0 \text{ or } A_j(z) \geq 0\}.$$

Choose $\tilde{z} \in H_{I_z}(0)$ such that $\|z - \tilde{z}\| = d(z, H_{I_z}(0))$. Note that by definition of I_z, $A_j(z) < 0$ for all $j \in (I_z)^c$. Consequently, $z \in \mathbb{B}_\varepsilon(\bar{z}) \cap H_{I_z}(\alpha)$ for α defined by

$$\alpha_j := A_j(z) \quad (j \in I_z).$$

Since also (6.9) ensures that $\alpha_i \in (-\varepsilon, \varepsilon)$ for all $i \in I_z$, we may apply (6.8) to derive that

$$d(z, H_{I_z}(0)) \leq L \|\alpha\| = L \sum_{j \in I_z} |A_j(z)|.$$

Now, if $j \in I_z$ is such that $A_j(z) + \vartheta_{1j} = 0$, then $|A_j(z)| = |\vartheta_{1j}|$. Otherwise, by (6.11), $A_j(z) + \vartheta_{1j} < 0$ and, by definition of I_z, $A_j(z) \geq 0$. This implies

$|A_j(z)| \le |\vartheta_{1j}|$. In any case we may conclude that

$$\|z - \tilde{z}\| = d(z, H_{I_z}(0)) \le L\,\|\vartheta_1\|\,.$$

Next, define $\tilde{\lambda} \in \mathbb{R}^p$ by $\tilde{\lambda}_j := \lambda_j + \vartheta_{2j}$ if $j \in I_z$ and $\tilde{\lambda}_j := 0$ if $j \in (I_z)^c$. Then, $\tilde{\lambda} \ge 0$ by (6.11). Moreover, $\tilde{z} \in H_{I_z}(0)$ entails that $A_j(\tilde{z}) = 0$ if $j \in I_z$ and $A_j(\tilde{z}) \le 0$ if $j \in (I_z)^c$. In particular, $\tilde{\lambda}_j A_j(\tilde{z}) = 0$ for all $j \in \{1, \ldots, p\}$. This means that $\tilde{\lambda} \in N_{\mathbb{R}^p_-}(A(\tilde{z}))$ and, hence, $(\tilde{z}, \tilde{\lambda}) \in M(0)$. Finally observe that, for $j \in (I_z)^c$, one has $A_j(z) + \vartheta_{1j} < 0$ and, thus, by (6.11), $\lambda_j = -\vartheta_{2j}$. This proves that $\tilde{\lambda} - \lambda = \vartheta_2$. Consequently,

$$\begin{aligned} d((z,\lambda), M(0)) &\le \|\,(z,\lambda) - (\tilde{z},\tilde{\lambda})\| = \|z - \tilde{z}\| + \|\lambda - \tilde{\lambda}\| \\ &\le L\|\vartheta_1\| + \|\vartheta_2\| \le (L+1)\,\|\vartheta\| \end{aligned}$$

which shows (6.10). □

In order to continue, we will need the following technical lemma.

Lemma 6.1. *Fix an arbitrary $I^* \subseteq \{1, \ldots, p\}$ and consider the following class of multifunctions*

$$\bar{H}_I(\alpha) := \{z | A_i(z) = \alpha_i \;\; (i \in I)\} \quad (I \subseteq \{1, \ldots, p\})$$

Assume that

1. *For all $I \neq I^*$ with $I^* \subseteq I \subseteq \{1, \ldots, p\}$ the $\bar{H}_I$ are calm at $(0, \bar{z})$.*
2. *For some $i' \in I \diagdown I^*$ the multifunctions*

$$\begin{aligned} M(\alpha,\beta) &:= \left\{ z \in \mathbb{R}^t \;\middle|\; \begin{array}{l} A_i(z) = \alpha_i \;\; (i \in I^*), \\ A_j(z) \le \beta_j \;\; (j \in \{1, \ldots, p\} \diagdown (I^* \cup \{i'\})) \end{array} \right\}, \\ \bar{M}(t) &:= \left\{ z \in \mathbb{R}^t \,| A_{i'}(z) = t \right\} \end{aligned}$$

   are calm at $(0, 0, \bar{z})$ and $(0, \bar{z})$, respectively.

Then, $\bar{H}_{I^*}$ is calm at $(0, \bar{z})$.

Proof. Assume that $\bar{H}_{I^*}$ fails to be calm at $(0, \bar{z})$. Then, by (2.1), there is a sequence $z_k \to \bar{z}$ such that

$$d(z_k, \bar{H}_{I^*}(0)) > k \left(\sum_{i \in I^*} |A_i(z_k)| + \sum_{j \in \{1,\ldots,p\} \diagdown I^*} [A_j(z_k)]_+ \right). \tag{6.12}$$

Suppose there is some index $j' \in \{1, \ldots, p\} \diagdown I^*$ and some subsequence z_{k_l} with $A_{j'}(z_{k_l}) \ge 0$. Put $I' := I^* \cup \{j'\}$. Due to $\bar{H}_{I'}(0) \subseteq \bar{H}_{I^*}(0)$ and to $z_{k_l} \in \bar{H}_{I'}(A(z_{k_l}))$ one would arrive from (6.12) at

$$d(z_{k_l}, \bar{H}_{I'}(0)) > k_l \left(\sum_{i \in I'} |A_i(z_{k_l})| + \sum_{j \in \{1,\ldots,p\} \diagdown I'} [A_j(z_{k_l})]_+ \right),$$

a contradiction with assumption 1. Hence, there is some k_0 such that

$$A_j(z_k) < 0 \quad \forall k \geq k_0 \quad \forall j \in \{1, \dots, p\} \diagdown I^*. \tag{6.13}$$

Together with (6.12), this implies that

$$d(z_k, \bar{H}_{I^*}(0)) > k \sum_{i \in I^*} |A_i(z_k)|. \tag{6.14}$$

We claim the existence of some $\rho > 0$ and $k_1 \geq k_0$ such that

$$\sum_{i \in I^*} |A_i(z_k)| > \rho |A_{i'}(z_k)| \quad \forall k \geq k_1, \tag{6.15}$$

where i' refers to assumption 2. Indeed, otherwise there is a subsequence z_{k_l} such that

$$\sum_{i \in I^*} |A_i(z_{k_l})| \leq l^{-1} |A_{i'}(z_{k_l})|.$$

In the following, we lead this relation to a contradiction. Now, justified by $\bar{z} \in \bar{M}(0) \neq \emptyset$, where $\bar{M}$ is defined in assumption 2, we may select for any l some $y_l \in \bar{M}(0)$ such that

$$d(z_{k_l}, \bar{M}(0)) = \|z_{k_l} - y_l\|.$$

The assumed calmness at $(0, \bar{z})$ of $\bar{M}$ entails the existence of some $L_1 > 0$ such that

$$d(z_{k_l}, \bar{M}(0)) \leq L_1 |A_{i'}(z_{k_l})| \to_l 0$$

which in turn implies that $y_l \to \bar{z}$. Consequently, for all large enough l,

$$|A_{i'}(z_{k_l})| = |A_{i'}(z_{k_l}) - A_{i'}(y_l)| \leq L_2 \|z_{k_l} - y_l\|$$

where L_2 is some Lipschitz modulus of $A_{i'}$ near $\bar{z}$. Now, referring to the multifunction M defined in assumption 2., we observe by virtue of (6.13) that, for all large enough l, $z_{k_l} \in M\left(\alpha^{(l)}, 0\right)$, where $\alpha_i^{(l)} := A_i(z_{k_l})$ for $i \in I^*$. Now, the assumed calmness at $(0, \bar{z})$ of M leads to

$$\begin{aligned} d(z_{k_l}, M(0,0)) \leq \quad & L_3 \left\|\alpha^{(l)}\right\| = L_3 \sum_{i \in I^*} |A_i(z_{k_l})| \leq l^{-1} L_3 |A_{i'}(z_{k_l})| \\ \leq \quad & l^{-1} L_3 L_2 \|z_{k_l} - y_l\| = l^{-1} L_3 L_2 d(z_{k_l}, \bar{M}(0)), \end{aligned}$$

for all large enough l. If also $l > L_3 L_2$, then

$$d(z_{k_l}, M(0,0)) < d(z_{k_l}, \bar{M}(0)). \tag{6.16}$$

Now, justified by $\bar{z} \in M(0,0) \neq \emptyset$, we may select $x_l \in M(0,0)$ such that

$$d(z_{k_l}, M(0,0)) = \|z_{k_l} - x_l\| \quad \forall l.$$

It follows from (6.16) that $x_l \notin \bar{M}(0)$, whence $A_{i'}(x_l) \neq 0$. Recalling that $A_{i'}(z_{k_l}) < 0$ for large enough l (see (6.13)), one would find in case of $A_{i'}(x_l) > 0$

some x' on the line segment $[z_{k_l}, x_l]$ with $A_{i'}(x') = 0$ and $\|z_{k_l} - x'\| < \|z_{k_l} - x_l\|$ yielding a contradiction with (6.16) due to $x' \in \bar{M}(0)$. Therefore, $A_{i'}(x_l) < 0$ and, hence, one may invoke the definition of M to infer from $x_l \in M(0,0)$ that $x_l \in \bar{H}_{I^*}(0)$ for large enough l. Now, (6.13) and (6.14) provide, for large enough l that

$$\begin{aligned} &k_l \left(\sum_{i \in I^*} |A_i(z_k)| + \sum_{j \in \{1,\dots,p\} \setminus (I^* \cup \{i'\})} [A_j(z_{k_l})]_+ \right) \\ &= k_l \sum_{i \in I^*} |A_i(z_{k_l})| < d(z_{k_l}, \bar{H}_{I^*}(0)) \le \|z_{k_l} - x_l\| = d(z_{k_l}, M(0,0)), \end{aligned}$$

a contradiction with the assumed calmness at $(0, 0, \bar{z})$ of M. This contradiction proves the desired relation (6.15). Using this, we may continue (6.14) as

$$\begin{aligned} d(z_k, \bar{H}_{I^*}(0)) &> k \left(\frac{1}{\rho+1} \sum_{i \in I^*} |A_i(z_k)| + \frac{\rho}{\rho+1} \sum_{i \in I^*} |A_i(z_k)| \right) \\ &> k \frac{\rho}{\rho+1} \left(\sum_{i \in I^* \cup \{i'\}} |A_i(z_k)| \right) \\ &= k \frac{\rho}{\rho+1} \left(\sum_{i \in I^* \cup \{i'\}} |A_i(z_k)| + \sum_{j \in \{1,\dots,p\} \setminus (I^* \cup \{i'\})} [A_j(z_k)]_+ \right) \\ &\forall k \ge k_1, \end{aligned}$$

where in the last relation, we exploited again (6.13). Put $I' := I^* \cup \{i'\}$. Due to $\bar{H}_{I'}(0) \subseteq \bar{H}_{I^*}(0)$ we end up at the relation

$$d(z_k, \bar{H}_{I'}(0)) > k \frac{\rho}{\rho+1} \left(\sum_{i \in I'} |A_i(z_k)| + \sum_{j \in \{1,\dots,m\} \setminus I'} [A_j(z_k)]_+ \right) \quad \forall k \ge k_1.$$

This, however, is in contradiction with the assumed calmness at $(0, \bar{z})$ of $\bar{H}_{I'}$ (see assumption 1.). Hence, we have finally led to a contradiction our initial assumption that $\bar{H}_{I^*}$ fails to be calm at $(0, \bar{z})$. □

We now show that the primal calmness condition can be reduced in such a way that only equality constrained subsystems need to be checked for calmness.

Proposition 6.3 (**reduction of the primal calmness condition**). *If for all $I \subseteq \{1, \dots, p\}$ the multifunctions*

$$\tilde{H}_I(\alpha) := \{z | A_i(z) = \alpha_i \ \ (i \in I)\}$$

are calm at $(0, \bar{z})$, then the multifunctions

$$\bar{H}_I(\alpha) = \{z | A_i(z) = \alpha_i \ \ (i \in I), \quad A_i(z) \le \alpha_i \ \ (i \in I^c)\}$$

are also calm at $(0, \bar{z})$ for all $I \subseteq \{1, \dots, p\}$. In particular, the multifunctions $H_I(\alpha)$ introduced in Proposition 6.2 are calm at $(0, \bar{z})$ for all $I \subseteq \{1, \dots, p\}$.

Proof. We proceed by induction over the number p of components of A. Consider first the case $p = 1$. We either have $I = \emptyset$ or $I = \{1\}$. In the second case, one has $\bar{H}_I = \tilde{H}_I$ due to $p = 1$, hence calmness of $\bar{H}_I$ follows from that of $\tilde{H}_I$. In the first case, we apply Lemma 6.1. Referring to the notation of this lemma, we put $I^* = \emptyset$ and check the two assumptions made there. As the only set $I \subseteq \{1\}$ with $I \neq I^*$ is given by $I = \{1\}$ and then, as before, $\bar{H}_I = \tilde{H}_I$, calmness of $\bar{H}_I$ follows from that of $\tilde{H}_I$. This shows the first assumption of Lemma 6.1 to hold true. Concerning the second assumption, one has $i' = 1$ and, hence, M reduces to the trivial constant multifunction $M(\alpha, \beta) \equiv \mathbb{R}^t$ which is calm. On the other hand, the second multifunction introduced there reduces to $\bar{M} = \tilde{H}_I$, hence calmness of $\bar{M}$ follows from that of $\tilde{H}_I$. As a consequence, Lemma 6.1 yields calmness of $\bar{H}_{I^*} = \bar{H}_\emptyset$. Summarizing, the assertion of our proposition follows for the case $p = 1$. Next assume that the Proposition holds true for all $p \leq k$ and consider the case $p = k + 1$. By assumption, the $\tilde{H}_I$ are calm at $(0, \bar{z})$ for all $I \subseteq \{1, \ldots, k+1\}$. In particular, the multifunction $\bar{M}$ considered in the second assumption of Lemma 6.1 and corresponding to the case $|I| = 1$ is calm at $(0, \bar{z})$. Moreover, the induction hypothesis ensures that also the multifunctions

$$\{z | A_j(z) = \alpha_j \ \ (j \in I), \quad A_j(z) \leq \alpha_j \ \ (j \in J \setminus I)\} \tag{6.17}$$

are calm at $(0, \bar{z})$ for all subsets $I \subseteq J$ and all $J \subseteq \{1, \ldots, k+1\}$ with $|J| = k$. Since the multifunction M considered in the second assumption of Lemma 6.1 is of type (6.17) with $J = \{1, \ldots, k+1\} \setminus \{i'\}$, it follows that M is calm at $(0, 0, \bar{z})$. Summarizing, the second assumption of Lemma 6.1 is always satisfied no matter how the index set $I^* \subseteq \{1, \ldots, k+1\}$ is chosen in the Lemma. Therefore it is enough to check the first assumption for its application.

Now, choose an arbitrary $I^* \subseteq \{1, \ldots, k+1\}$. We have to show that $\bar{H}_{I^*}$ is calm at $(0, \bar{z})$. If $I^* = \{1, \ldots, k+1\}$, then $\bar{H}_{I^*} = \tilde{H}_{I^*}$ and calmness of $\bar{H}_{I^*}$ follows from that of $\tilde{H}_{I^*}$. If $|I|^* = k$, then the only choice for the index set I considered in the first assumption of Lemma 6.1 is $I = \{1, \ldots, k+1\}$. According to what we have shown just before, $\bar{H}_I$ is calm, so we have shown that the $\bar{H}_{I^*}$ are calm at $(0, \bar{z})$ whenever $|I^*| \geq k$. Passing to the case $|I^*| = k - 1$ and recalling that the index set I considered in the first assumption of Lemma 6.1 is always strictly larger than I^*, one derives calmness of $\bar{H}_I$ on the basis of what we have shown before due to $|I| > |I^*| = k - 1$ which amounts to $|I| \geq k$. So, the first assumption of Lemma 6.1 is satisfied again and we derive calmness of $\bar{H}_{I^*}$ whenever $|I^*| \geq k - 1$. Proceeding this way until $|I^*| = 0$, we get the desired calmness at $(0, \bar{z})$ for all subsets $I^* \subseteq \{1, \ldots, k+1\}$.

That the calmness of the $\bar{H}_I$ implies the calmness of the corresponding H_I introduced in Proposition 6.2, is an immediate consequence of the calmness definition and of the evident relations $\bar{H}_I(\alpha, 0) = H_I(\alpha)$. □

Using Propositions 6.2 and 6.3 along with Theorem 6.3, we can provide an assumption which completely relies on constraint systems induced by A and thus can be considered to be a CQ (weaker than surjectivity) for the mapping A.

Theorem 6.4 (**upper-approximation using MFCQ and CQ***). *In the setting of Theorem 6.3 assume that*

1. *MFCQ is satisfied at* $\bar{z} \in C$*;*

2. *all perturbed equality subsystems*

$$\{z \mid A_i(z) = \alpha_i \ (i \in I)\} \quad I \subseteq \{1, \dots, p\} \tag{6.18}$$

   are calm at $(0, \bar{z})$.

Then the coderivative formula of Theorem 6.3 holds true. We refer to (6.18) as "CQ**".*

Though Assumption 2. in Theorem 6.4 is much simpler than the assumption it replaced in Theorem 6.3, one still needs a concrete way of verifying calmness. This fact highlights another reason for the rising interest in providing ways to check for calmness of multifunctions. We refer to the beginning of Chapter 5, where literature pertaining to verification of calmness was briefly discussed.

Apart from verifying CQ* with the results just mentioned, we present in the next result another constraint qualification whose fulfillment guarantees Assumption 2. in Theorem 6.4 holds.

Proposition 6.4 (**a constraint qualification ensuring calmness**). *Assume that at* $\bar{z}$ *the following full rank constraint qualification is satisfied:*

$$\textit{rank } \{\nabla A_j(\bar{z})\}_{j \in I} = \min\{t, |I|\} \quad \forall I \subseteq \{1, \dots, p\}. \tag{6.19}$$

Then, the multifunctions $\tilde{H}_I$ *introduced in Proposition 6.3 are calm at* $(0, \bar{z})$ *for all* $I \subseteq \{1, \dots, p\}$.

Proof. Choose an arbitrary $I \subseteq \{1, \dots, p\}$. Consider first the case that $|I| \leq n$. Then, by (6.19), the set of gradients $\{\nabla A_j(\bar{z})\}_{i \in I}$ is linearly independent. Consequently, $\tilde{H}_I$ is calm at $(0, \bar{z})$. Now, if $|I| > n$, then select an arbitrary $J \subseteq I$ with $|J| = n$. By (6.19), the set of gradients $\{\nabla A_j(\bar{z})\}_{i \in J}$ is linearly independent, hence $\tilde{H}_J(0) = \{\bar{z}\}$ by the inverse function theorem. Since $A(\bar{z}) = 0$ and $\tilde{H}_I(0) \subseteq \tilde{H}_J(0)$, it follows that $\tilde{H}_I(0) = \tilde{H}_J(0)$. Moreover, according to what has been mentioned before, $\tilde{H}_J$ is calm at $(0, \bar{z})$. Consequently, there are constants $L, \varepsilon > 0$ such that

$$d(z, \tilde{H}_J(0)) \leq L \|\tilde{\alpha}\| \quad \forall z \in \tilde{H}_J(\tilde{\alpha}) \cap \mathbb{B}_\varepsilon(\bar{z}) \quad \forall \tilde{\alpha} \in \mathbb{B}_\varepsilon(0).$$

$>$From here it follows with $\tilde{H}_I(\alpha) \subseteq \tilde{H}_J(\tilde{\alpha})$, where $\tilde{\alpha}$ is the subvector of α according to the index set $J \subseteq I$, that

$$d(z, \tilde{H}_I(0)) = d(z, \tilde{H}_J(0)) \leq L \|\tilde{\alpha}\| \leq L \|\alpha\| \quad \begin{array}{l} \forall z \in \tilde{H}_I(\alpha) \cap \mathbb{B}_\varepsilon(\bar{z}) \\ \forall \alpha \in \mathbb{B}_\varepsilon(0) \end{array}.$$

This, however, amounts to calmness of $\tilde{H}_I$ at $(0, \bar{z})$. $\square$

Next, we demonstrate how CRCQ can also be used to verify CQ*.

Lemma 6.2 (CRCQ $\Rightarrow$ CQ*). *Let $\bar{z} \in C$ such that $A(\bar{z}) = 0$ and assume that CRCQ holds at $\bar{z}$. Then the multifunctions $\tilde{H}_I$ introduced in Proposition 6.3 are calm at $(0, \bar{z})$ for all $I \subseteq \{1, \ldots, p\}$.*

Proof. Define $\widetilde{C} := \{z \in \mathbb{R}^t \,|A_j(z) = 0, j = 1, \ldots, p\}$. As CRCQ condition concerns active constraints, regardless of whether they are inequalities or equalities, it is clear that CRCQ holds at $\bar{z} \in \widetilde{C}$. Given $A(\bar{z}) = 0$, it is easy to see that $\widetilde{C} \subseteq \tilde{H}_I(0)$, and therefore $\bar{z} \in \tilde{H}_I(0)$ for all subsets $I \subseteq \{1, \ldots, p\}$.

By the assumption, there exists of a neighborhood $\mathcal{U}$ of $\bar{z}$ such that the rank $\operatorname{rank}\{\nabla A_J(z)\}$ remains constant on $\mathcal{U}$ for each $J \subseteq \{1, \ldots, p\}$, thus it is easy to see that CRCQ must hold at $\bar{z} \in H_I(0)$ for all subsets $I \subseteq \{1, \ldots, p\}$. Then by Theorem 3 in Minchenko and Stakhovski [2009] and Theorem 3.1 in Song [2006] each H_I is calm at $(0, \bar{z})$. □

Remark 6.1. *The key to the previous proof relies on the sets $H_I(0)$ satisfying CRCQ at $\bar{z}$. In such a situation, these sets also satisfy the so-called* relaxed CRCQ *at $\bar{z}$ (Minchenko and Stakhovski [2009, Definition 1]) which states: For a set $C' := \{z \in \mathbb{R}^t \,|A_j(z) \leq 0, j \in K, A_j(z) = 0, j \in K_0\}$, where $K \cup K_0 = \{1, \ldots, p\}$, the relaxed CRCQ holds at $\bar{z} \in C'$ if there exists a neighborhood $\mathcal{V}$ of $\bar{z}$ such that* $\operatorname{rank}\{\nabla A_J(z)\}$ *remains constant on $\mathcal{V}$ for all $J = K' \cup K_0$, where $K' \subseteq I(\bar{z})$, i.e., the active index set.*

Given Lemma 6.2, we can provide the following corollary to Theorem 6.4

Corollary 6.2 (an upper-approximation using MFCQ and CRCQ). *Let $\bar{z} \in C$ such that $A(\bar{z}) = 0$ and assume that the following conditions hold*

1. *MFCQ at $\bar{z} \in C$*

2. *CRCQ at $\bar{z} \in C$*

Then the upper estimate of $D^* N_C$ in Theorem 6.3 holds

Corollary 6.2 places the conditions for the upper approxmation back into the familiar jargon of non-linear programming. At this point, we have all but exhausted the possibilities of the regular and non-regular setting using standard constraint qualifications and calmness. We will now see that calmness plays a crucial role in the ability to obtain explicit transformation formulae.

6.3 Beyond Calmness

In this final section, we briefly address some of the issues concerning the non-linear case when only MFCQ is known to hold at the point in question. Consider now the following example (Henrion et al. [2009c, Example 3.6])

Example 6.2. *Let* $A(z_1, z_2) := (-z_2, \varphi(z_1) - z_2)$*, where* $\varphi(t) := t^5 \sin(1/t)$ *for* $t \neq 0$ *and* $\varphi(0) := 0$*. Since* φ *is twice continuously differentiable, so is* A *and one has*

$$\nabla A_1(z_1, z_2) = (0, -1), \quad \nabla A_2(z_1, z_2) = (\varphi'(z_1), -1). \tag{6.20}$$

We choose $\bar{z} := (0, 0)$*. Then,* $A(\bar{z}) = (0, 0)$ *and, taking into account that* $\varphi'(0) = 0$*, it holds that*

$$\nabla A_1(\bar{z}) = \nabla A_2(\bar{z}) = (0, -1). \tag{6.21}$$

This means that both gradients are positively linearly independent (i.e., MFCQ is satisfied). Of course they are linearly dependent, thus preventing us from applying (6.7). Summarizing, all assumptions of Theorem 6.3 are fulfilled except calmness (this could be easily checked directly, but we shall see it as a consequence of the conclusion of that theorem being violated). We choose $v^* := 0$ *and* $\bar{v} := \nabla A_1(\bar{z}) + \nabla A_2(\bar{z})$ *(implying that* $\bar{v} \in N_C(\bar{z})$ *in view of MFCQ). Formally applying Theorem 6.3 would yield the inclusion*

$$D^* N_C(\bar{z}, \bar{v})(0) \subseteq \bigcup_{\bar{\lambda} \geq 0, \nabla^T A(\bar{z}) = \bar{v}} \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^2_-}\left(0, \bar{\lambda}\right)(0).$$

Given the fact that $D^* N_{\mathbb{R}^2_-}\left(0, \bar{\lambda}\right)(0) = \mathbb{R}^2$*, regardless of the value of* $\bar{\lambda} \geq 0$ *(see Proposition 6.1), and taking into account (6.21), we end up at the inclusion*

$$D^* N_C(\bar{z}, \bar{v})(0) \subseteq \{0\} \times \mathbb{R}. \tag{6.22}$$

To see that this is wrong, consider the sequences

$$z^k := (1/(k\pi), 0), \quad v_k := \nabla A_1\left(z^k\right) + \nabla A_2\left(z^k\right).$$

Then taking into account that $\varphi(z_1^k) = 0$ *and, thus,* $A\left(z^k\right) = (0, 0)$*, it follows that*

$$z^k \to \bar{z}, \quad z^k \in C, \quad v_k \to \bar{v}, \quad v_k \in N_C\left(z^k\right).$$

Here, the last relation relies on the fact that MFCQ is an open property and and pertains to hold at z^k *close to* $\bar{z}$*. Furtherore, we observe that* $\varphi'(z_1^k) \neq 0$*, which, as a consequence of (6.20), implies* $\nabla A\left(z^k\right)$ *is surjective; in fact,* $\nabla A\left(z^k\right)$ *is even a regular matrix. This allows to apply (6.7) at* $\left(z^k, v_k\right)$*:*

$$D^* N_C\left(z^k, v_k\right)(0) = \nabla^T A\left(z^k\right) D^* N_{\mathbb{R}^2_-}(0, (1, 1))(0) = \nabla^T A\left(z^k\right) \mathbb{R}^2 = \mathbb{R}^2,$$

where the last equality follows from the fact that,

$$\text{rank}\, \nabla^T A\left(z^k\right) = \text{rank}\, \nabla A\left(z^k\right) = 2.$$

Exploiting the robustness property of the co-derivative, we let $k \to \infty$ *and thus derive*

$$\mathbb{R}^2 \supseteq D^* N_C(\bar{z}, \bar{v})(0) \supseteq \operatorname*{Lim\,sup}_{k \to \infty} D^* N_C\left(z^k, v_k\right)(0) = \mathbb{R}^2,$$

therefore $D^* N_C(\bar{z}, \bar{v})(0) = \mathbb{R}^2$, *contradicting (6.22).*

We see then that by dropping the calmness condition, one can no longer expect the formula of Theorem 6.3 to hold true. Nevertheless, the formula may serve as a part of calculating the co-derivative in a more elementary (according to its basic definition) aggregation process. More precisely, by introducing the set

$$C^* := \{z \in \operatorname{bd} C | \nabla A(z) \text{ is surjective}\},$$

we have for any $\bar{v}^* \in \mathbb{R}^n$ (see Chapter 2),

$$D^* N_C(\bar{z}, \bar{v})(\bar{v}^*) = \operatorname*{Lim\,sup}_{\substack{(z,v,v^*)\to(\bar{z},\bar{v},\bar{v}^*)\\ z\in C\\ v\in N_C(z)}} \hat{D}^* N_C(z, v)(v^*),$$

which leads to the following expression

$$\begin{aligned}\operatorname*{Lim\,sup}_{\substack{(z,v)\to(\bar{z},\bar{v})\\ v\to\bar{v}^*\\ z\in C\\ v\in N_C(z)}} \hat{D}^* N_C(z, v)(v^*) &= \operatorname*{Lim\,sup}_{\substack{(z,v)\to(\bar{z},\bar{v})\\ v\to\bar{v}^*\\ z\in \operatorname{int} C\\ v\in N_C(z)}} \hat{D}^* N_C(z, v)(v^*) \cup \\ &\quad \underbrace{\operatorname*{Lim\,sup}_{\substack{(z,v)\to(\bar{z},\bar{v})\\ v\to\bar{v}^*\\ z\in C\setminus C^*\\ v\in N_C(z)}} \hat{D}^* N_C(z, v)(v^*)}_{P(\bar{v}^*)} \cup \underbrace{\operatorname*{Lim\,sup}_{\substack{(z,v)\to(\bar{z},\bar{v})\\ v\to\bar{v}^*\\ z\in C^*\\ v\in N_C(z)}} D^* N_C(z, v)(v^*)}_{R(\bar{v}^*)} \\ &= \{0\} \cup P(\bar{v}^*) \cup R(\bar{v}^*)\end{aligned}$$

Here, the first term is trivial and follows easily from the definition of the Fréchet coderivative evaluated in the interior of the feasible set, whereas the last equality results from the robustness (outer semicontinuity) of the co-derivative. In this way, we have subdivided the computation of the co-derivative into a **p**athological part $P(\bar{v}^*)$, where Fréchet coderivatives have to be calculated and aggregated in an elementary way, and a **r**egular part $R(\bar{v}^*)$, where we may exploit formula (6.7) in the aggregation process. It is important to observe that the pathological part is small in the following sense (Henrion [1992, Theorem 2.1]): If the MFCQ is satisfied everywhere in the feasible set C, then the subset of points around which the feasible set may be locally described by a regular constraint system (i.e., with surjective Jacobian $\nabla A(z)$) is open and dense in the boundary of C.

It should now be obvious that it will be rather difficult to obtain useful transformation formulae or upper approximations of $D^* N_C$ in the absence of calmness. Nevertheless, Henrion et al. [2009c, Proposition 3.4, 3.5, and Corollary 3.4] provide some hints as to how the regular component $R(\bar{v}^*)$ may look. Due to the fact that in the applications we will not encounter such situations, we leave out these last few results.

Chapter 7

Structural Properties of the Spot Market EPEC

In this chapter, we determine structural properties associated with the spot market EPEC (3.9) outlined in Chapter 3 that are needed for the derivation of explicit stationarity conditions and a solution analysis. In order to unburden the analysis from technical difficulties associated with rare cases, we introduce a restricted class of solutions to (3.9). More precisely, we consider solutions $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ such that

$$\boxed{\begin{array}{cc} \bar{\alpha}_i, \bar{\beta}_i > 0 & i = 1, \ldots, N \\ \bar{q}_i > 0 & i = 1, \ldots, l \\ \bar{q}_i = 0 & i = l+1, \ldots, N \\ -\hat{y}_j < \bar{y}_j < \hat{y}_j & j = 1, \ldots, k \\ \bar{y}_j = \hat{y}_j & j = k+1, \ldots, m \end{array}} \tag{7.1}$$

As in Henrion and Römisch [2007], we only consider strictly positive bidding coefficients. Disregarding zero coefficients allows us to avoid economically nonsensical or more pathological situations. In addition to the types of solutions analyzed in Henrion and Römisch [2007], we consider cases wherein certain transmission lines become congested, electricity is lost due to resistance, and some producers may not be participating at equilibrium. These considerations add a new level of difficulty to the analysis. More precisely, we assume without loss of generality that only the first $l \geq 1$ generators are active[1] and that the first k transmission lines are uncongested.

As mentioned in the introduction, the identification of situations where certain market participants are forced to become non-active is of economic, yet at the same time, mathematical interest. Indeed, it is here where the nonsmooth character of EPECs emerges. We note then, that it would only add to notational, not mathematical, difficulty to consider generation quantities reaching their respective upper bounds. Similarly, when $m - k > 0$, i.e., some of the transmission lines are congested, we assume only the upper transmission capacity has been reached.

[1] note that $l = 0$ is excluded as we have assumed that total demand is positive, thus implying that at least one producer is active at a solution

7.1 The ISO problem

In this section, we compile certain structural properties of the ISO problem (3.7). For convenience, we partition $q = (q^{(1)}, q^{(2)}) \in \mathbb{R}^l \times \mathbb{R}^{N-l}$ and $y = (y^{(1)}, y^{(2)}) \in \mathbb{R}^k \times \mathbb{R}^{m-k}$. Given the feasible set G from the ISO-Problem (3.7) it is easy to see that near solutions satisfying (7.1), G can be described by

$$G = \left\{(q, y) \in \mathbb{R}^{N+m} \,|H(q, y) \leq 0\right\}, \tag{7.2}$$

where $H : \mathbb{R}^{N+m} \to \mathbb{R}^{2N-l+m-k}$ is the twice continuously differentiable mapping defined by the inequalities that are active near solutions of the type given in (7.1), i.e.,

$$H(q, y) := \begin{pmatrix} d + L(y) - q - By \\ -q^{(2)} \\ y^{(2)} - \hat{y}^{(2)} \end{pmatrix}. \tag{7.3}$$

In what follows, we will need some auxiliary statements.

Lemma 7.1 (properties of the incidence matrix). *Let B be any (N, m)-incidence matrix of some oriented connected graph. Then the following properties hold (with $\hat{y}$ referring to the vector of upper transmission bounds):*

1. $\ker B^T = \mathbb{R}(1, \ldots, 1)^T$
2. *For any integer p such that $1 \leq p \leq N$, each $(N-p, m)$-submatrix of B has rank $N - p$.*
3. $\forall \varepsilon > 0 \quad \exists \Delta' > 0 \quad \forall \rho_j \in [0, \Delta') \quad \forall y \in [-\hat{y}, \hat{y}] : \quad \|\nabla L(y)\| < \varepsilon.$
4. $\exists \Delta'' > 0 \quad \forall \rho_j \in [0, \Delta'') \quad \forall y \in [-\hat{y}, \hat{y}]$:

$$\textit{if } \nabla^T L(y)z = B^T z \textit{ and } z_i = 0 \textit{ for some } i \textit{ then } z = 0.$$

Proof. For 1., see Biggs [1994, Proposition 4.3]. In particular, $\operatorname{rank} B = N-1$. For 2., assume that the rank of some $(N-p, m)$-submatrix of B is smaller than $N-p$. Then by successively joining the p rows removed from B to this submatrix and thus, reconstructing B, the rank can increase by at most $p-1$, since the last row is already a linear combination of all the remaining $N-1$ rows (see 1.). Whence a contradiction with $\operatorname{rank} B = N - 1$ (see 1.). 3. is an immediate consequence of (3.5), i.e., continuity. Concerning 4., it follows from 3. that for small enough transmission losses, $\nabla L(y)$ can be considered arbitrarily small for all y in the indicated compact range. Since $\operatorname{rank} B = N - 1$, one has $\operatorname{rank}(\nabla L(y) - B) \geq N - 1$ for small losses. If this rank strictly increases, then the dimension of the corresponding kernel strictly decreases, hence $\ker(\nabla^T L(y) - B^T) = \{0\}$ by 1. Otherwise, this rank remains $N - 1$, hence the corresponding kernel stays one-dimensional. Now by 1. and a continuity argument there exists some $v \neq 0$, which can be chosen arbitrarily close to $(1, \ldots, 1)^T$, such that $\ker(\nabla^T L(y) - B^T) = \mathbb{R}v$. In either case, the asserted implication in 4. follows. $\square$

The following lemma provides some initial properties of the constraint mapping H defined in (7.3). In particular, it clarifies under which conditions the inequality system (7.2) satisfies MFCQ or LICQ. Moreover, information pertaining to the Lagrange multipliers associated with the ISO problem (3.7) is derived. To do so, we will need to split the matrix $\nabla L(y) - B$ into specific submatrices according to the parameters l and k of activity and congestion:

$$\nabla L(y) - B = \begin{pmatrix} \nabla L^1(y) - B^1 \\ \nabla L^2(y) - B^2 \end{pmatrix} = \begin{pmatrix} \nabla L^{11}(y) - B^{11} & \nabla L^{12}(y) - B^{12} \\ \nabla L^{21}(y) - B^{21} & \nabla L^{22}(y) - B^{22} \end{pmatrix}.$$

Lemma 7.2 (LICQ/MFCQ, Constraint Activity, and Complementarity). *Let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution to (3.9) satisfying (7.1). Then there exists some $\Delta > 0$ such that under the condition $\rho_j \in [0, \Delta)$ $(j = 1, \ldots, m)$ the following properties hold:*

1. a) *The rows of $\nabla H(\bar{q}, \bar{y})$ are positively linearly independent. (MFCQ)*
   b) *If $l = N$ (all generators active) or $k = m$ (no congestion) then $\nabla H(\bar{q}, \bar{y})$ is surjective. (LICQ)*
   c) *If $1 \leq l < N$, $0 \leq k < m$, and B^{21} is surjective, then $\nabla H(\bar{q}, \bar{y})$ is surjective. (LICQ)*
2. *$H(\bar{q}, \bar{y}) = 0$.*
3. a) *Concerning the ISO problem (3.7) (with fixed parameters $(\bar{\alpha}, \bar{\beta})$), strict complementarity holds for the first l demand-satisfaction constraints (3.6) (corresponding to active producers).*
   b) *If $k = m$ (no congestion) or all multipliers associated with the flow constraints $y^{(2)} \leq \hat{y}^{(2)}$ vanish, then strict complementarity holds for all demand-satisfaction constraints (3.6).*

Proof. The Jacobian of the mapping $H(q, y)$ at $(\bar{q}, \bar{y})$ becomes

$$\nabla H(\bar{q}, \bar{y}) = \begin{pmatrix} -I_1 & 0 & \nabla L^{11}(\bar{y}) - B^{11} & \nabla L^{12}(\bar{y}) - B^{12} \\ 0 & -I_2 & \nabla L^{21}(\bar{y}) - B^{21} & \nabla L^{22}(\bar{y}) - B^{22} \\ 0 & -I_2 & 0 & 0 \\ 0 & 0 & 0 & I_3 \end{pmatrix}. \tag{7.4}$$

Assume a relation $\nabla^T H(\bar{q}, \bar{y})c = 0$ for some $c \geq 0$.

$$\begin{pmatrix} -I_1 & 0 & 0 & 0 \\ 0 & -I_2 & -I_2 & 0 \\ (\nabla L^{11}(\bar{y}) - B^{11})^T & (\nabla L^{21}(\bar{y}) - B^{21})^T & 0 & 0 \\ (\nabla L^{12}(\bar{y}) - B^{12})^T & (\nabla L^{22}(\bar{y}) - B^{22})^T & 0 & I_3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0.$$

Then

$$c_1 = 0,\ c_2 = -c_3,\ (\nabla L^{21}(\bar{y}) - B^{21})^T c_2 = 0,\ (\nabla L^{22}(\bar{y}) - B^{22})^T c_2 = -c_4. \tag{7.5}$$

Given that $c_2, c_3 \geq 0$, it follows that $c_2 = c_3 = 0$, and so $c_4 = 0$ as well. Therefore, $c = 0$, which proves the positive linear independence of the rows of $\nabla H(\bar{q}, \bar{y})$ as claimed in 1. (a). Inspecting again (7.5), we observe, that the conclusion $c = 0$ could equally well be drawn from the relation $\nabla^T H(\bar{q}, \bar{y})c = 0$ upon replacing the original assumption $c \geq 0$ by the injectivity of $(\nabla L^{21}(\bar{y}) - B^{21})^T$. This, however, follows from the assumed surjectivity of B^{21} if $\Delta > 0$ in the statement of our lemma is chosen small enough as to maintain surjectivity of $\nabla L^{21}(\bar{y}) - B^{21}$ via statement 3. of Lemma 7.1. This shows 1. (c). Concerning 1. (b), consider first the case where $l = N$. Here, the third row block of $\nabla H(\bar{q}, \bar{y})$ in (7.4) is missing, in which case surjectivity of $\nabla H(\bar{q}, \bar{y})$ is obvious. If instead $k = m$, then the fourth row block of $\nabla H(\bar{q}, \bar{y})$ in (7.4) is missing. Then surjectivity of $\nabla H(\bar{q}, \bar{y})$ follows via the surjectivity of $(\nabla L^2(\bar{y}) - B^2)$. Indeed, Lemma 7.1 (statement 2.) states that B^2 has rank $N - l$. Then as a consequence of Lemma 7.1 (statement 3.), there exists some $\Delta > 0$ such that $(\nabla L^2(\bar{y}) - B^2)$ has rank $N - l$ too, provided $\rho_j \in [0, \Delta)$, for all $j = 1, \ldots, m$ and $\Delta > 0$. In other words, $(\nabla L^2(\bar{y}) - B^2)$ is surjective and 1. (b) is proven.

Statement 1.(a) guarantees the existence of Lagrange multipliers such that the first-order optimality conditions of (3.7) hold for a solution $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$. Accordingly, there exist $\bar{\lambda}^1, \bar{\lambda}^2, \bar{\mu}, \bar{\eta} \geq 0$ such that

$$\begin{pmatrix} \bar{\alpha} + 2[\mathrm{diag}\ \bar{\beta}]\bar{q} \\ 0 \end{pmatrix} + \nabla^T H(\bar{q}, \bar{y}) \begin{pmatrix} \bar{\lambda}^{(1)} \\ \bar{\lambda}^{(2)} \\ \bar{\mu} \\ \bar{\eta} \end{pmatrix} = 0 \tag{7.6}$$

$$H(\bar{q}, \bar{y}) \bullet (\bar{\lambda}^1, \bar{\lambda}^2, \bar{\mu}, \bar{\eta}) = 0 \tag{7.7}$$

Here, we let '$\bullet$' denote the Hadamard, or component-wise, product. Then by taking into account (7.1) we obtain the following set of relations

$$0 < \bar{\alpha}_i + 2\bar{\beta}_i \bar{q}_i = \bar{\lambda}_i^{(1)}, \quad i = 1, \ldots, l \tag{7.8}$$

$$0 < \bar{\alpha}_i = \bar{\lambda}_{i-l}^{(2)} + \bar{\mu}_{i-l}, \quad i = l+1, \ldots, N \tag{7.9}$$

$$0 = (\nabla L^{11}(\bar{y}) - B^{11})^T \bar{\lambda}^{(1)} + (\nabla L^{21}(\bar{y}) - B^{21})^T \bar{\lambda}^{(2)} \tag{7.10}$$

$$-\eta = (\nabla L^{12}(y) - B^{12})^T \bar{\lambda}^{(1)} + (\nabla L^{22}(y) - B^{22})^T \bar{\lambda}^{(2)}. \tag{7.11}$$

From (7.8) we derive statement 3. (a). Under any of the assumptions of statement 3. (b), (7.10) and (7.11) combine to

$$(\nabla L(\bar{y}) - B)^T \begin{pmatrix} \bar{\lambda}^{(1)} \\ \bar{\lambda}^{(2)} \end{pmatrix} = 0.$$

Then by choosing Δ in the assertion of this lemma equal to Δ'' in statement 4. of Lemma 7.1, we may draw the following conclusion: if there exists $i \in \{l+1, \ldots, N\}$ such that $\bar{\lambda}_{i-l}^{(2)} = 0$, then $\bar{\lambda}^{(1)} = \bar{\lambda}^{(2)} = 0$, contradicting (7.8). Therefore, $\bar{\lambda}_{i-l}^{(2)} > 0$ and statement 3. (b) follows.

In order to prove statement 2., we first observe that $H_i(\bar{q}, \bar{y}) = 0$ for $i > N$

because by definition of $q^{(2)}$ and $y^{(2)}$ in (7.3) via (7.1), one has that $\bar{q}^{(2)} = 0$ and $\bar{y}^{(2)} = \hat{y}^{(2)}$. Moreover, (7.8) along with (7.7) implies $H_i(\bar{q}, \bar{y}) = 0$ for $i \leq l$. It remains to prove that $H_i(\bar{q}, \bar{y}) = 0$ for $i \in \{l+1, \ldots, N\}$. This would follow easily from (7.7) and statement 3. (b) of this lemma in the case of $k = m$ (no congestion), however, the case of congestion requires a more subtle reasoning. Therefore, assume there exists some $i^* \in \{l+1, \ldots, N\}$ such that $H_{i^*}(\bar{q}, \bar{y}) < 0$. Given the structure of H and noting that the loss function L has nonnegative components, this amounts to saying that

$$0 \leq d_{i^*} + L_{i^*}(\bar{y}) < \bar{q}_{i^*} + \sum_{j=1}^{m} b_{i^*j}\bar{y}_j. \tag{7.12}$$

Recalling that i^* is chosen among the set of inactive producers, we derive that

$$\sum_{j=1}^{m} b_{i^*j}\bar{y}_j > 0. \tag{7.13}$$

Now, for any $i^1 \in \{1, \ldots, N\}$ we declare $i^2 \in \{1, \ldots, N\}$ to be a *critical neighbor* of i^1 if there exists some $j \in \{1, \ldots, m\}$ such that

$$|b_{i^1j}| = |b_{i^2j}| = 1,\ b_{i^1j}b_{i^2j} = -1,\ b_{i^1j}\bar{y}_j > 0. \tag{7.14}$$

We define the set $\Pi \subseteq \{1, \ldots, N\}$ to consist of our fixed node i^*, all its critical neighbors, all the critical neighbors of these critical neighbors etc. Define $\Lambda \subseteq \{1, \ldots, m\}$ to consist of those transmission lines connecting nodes inside Π only. Then, $\Pi \times \Lambda$ constitutes a subgraph of the original one which is connected again by construction (there exists a path from any node in Π to i^*). Consequently, the associated submatrix of B is again the incidence matrix of a connected oriented graph. Therefore, we can invoke statement 1. of Lemma 7.1 to conclude that $\sum_{i\in\Pi} b_{ij} = 0$ for all $j \in \Lambda$. This in turn implies that

$$\sum_{i\in\Pi}\sum_{j\in\Lambda} b_{ij}\bar{y}_j = 0. \tag{7.15}$$

Next we observe that

$$\sum_{j\in\Lambda^c} b_{ij}\bar{y}_j \leq 0 \quad \forall i \in \Pi. \tag{7.16}$$

Indeed, otherwise there exists some $i \in \Pi$ and $j \in \Lambda^c$ with $b_{ij}\bar{y}_j > 0$. In particular, $b_{ij} \neq 0$, whence $|b_{ij}| = 1$. Moreover, let $i^a \in \{1, \ldots, N\}$ be the uniquely defined node such that $b_{ij}b_{i^aj} = -1$ (i.e., i^a is the node joined with i via edge j). Then, by definition, i^a is a critical neighbor of i, whence $i^a \in \Pi$. Therefore, the edge j joining i and i^a belongs to Λ which contradicts $j \in \Lambda^c$. Now, combining (7.13) with (7.16) yields $\sum_{j\in\Lambda} b_{i^*j}\bar{y}_j > 0$ which along with $i^* \in \Pi$ and (7.15) allows to infer the existence of some $i^{**} \in \Pi$ such that $\sum_{j\in\Lambda} b_{i^{**}j}\bar{y}_j < 0$. Then, the demand satisfaction at i^{**} provides that (taking into account (7.16) for i^{**})

$$0 \leq d_{i^{**}} + L_{i^{**}}(\bar{y}) \leq \bar{q}_{i^{**}} + \sum_{j=1}^{m} b_{i^{**}j}\bar{y}_j < \bar{q}_{i^{**}}.$$

In other words, i^{**} is an active producer, hence $i^{**} \leq l$. Finally, we observe, that for each critical neighbor i' of i^* we may modify the flow vector $\bar{y}$ to some $\tilde{y}$ such that $(\bar{q}, \tilde{y})$ remains feasible for the ISO problem (3.7) (i.e., $H(\bar{q}, \tilde{y}) \leq 0$) and that $H_{i'}(\bar{q}, \tilde{y}) < 0$. Indeed, assuming that i' and i^* are joined by some edge j', we may define $\tilde{y}$ as

$$\tilde{y}_k := \begin{cases} \bar{y}_k & k \neq j' \\ \bar{y}_{j'} - \varepsilon & k = j', b_{i^*j'} = 1 \\ \bar{y}_{j'} + \varepsilon & k = j', b_{i^*j'} = -1 \end{cases},$$

where $\varepsilon > 0$ is chosen small enough to guarantee that the demand at node i^* remains satisfied after the modification (which is possible by (7.12)):

$$\bar{q}_{i^*} + \sum_{j=1}^{m} b_{i^*j}\tilde{y}_j \geq d_{i^*} + L_{i^*}(\bar{y}). \tag{7.17}$$

Moreover, let $\varepsilon > 0$ be small enough such that $\tilde{y}_j \leq \hat{y}_j$. This is possible due to $\bar{y}_j \leq \hat{y}_j$ (by feasibility of $\bar{y}$) and upon observing that $\bar{y}_j > 0$ in case of $b_{i^*j} = 1$ and $\bar{y}_j < 0$ in case of $b_{i^*j} = -1$ (see (7.14)). Thus, $\tilde{y}$ is feasible for the constraint $y \leq \hat{y}$. In addition, we have that $b_{i'j'}\tilde{y}_{j'} = b_{i'j'}\bar{y}_{j'} + \varepsilon$ by construction of $\tilde{y}_{j'}$ and by $b_{i^*j'} = -b_{i'j'}$. Then, the demand satisfaction at node i' reads as

$$d_{i'} + L_{i'}(\bar{y}) \leq \bar{q}_{i'} + \sum_{j=1}^{m} b_{i'j}\bar{y}_j = \bar{q}_{i'} + \sum_{j=1}^{m} b_{i'j}\tilde{y}_j - \varepsilon.$$

Since the demand satisfaction relations at nodes i' and i^* are the only ones affected by the transition from $\bar{y}$ to $\tilde{y}$, it follows that $(\bar{q}, \tilde{y})$ remains feasible for (3.7) (i.e., $H(\bar{q}, \tilde{y}) \leq 0$) and $H_{i'}(\bar{q}, \tilde{y}) < 0$ (demand is strictly exceeded by offer). In this way, we have shifted the strict inequality $H_{i^*}(\bar{q}, \bar{y}) < 0$ to the strict inequality $H_{i'}(\bar{q}, \tilde{y}) < 0$ at any of the critical neighbors i' of i^* just by modifying the flow vector. This procedure can now be repeated for any of the critical neighbors of i', and so after finitely many steps one arrives at a feasible solution $(\bar{q}, y')$ of (3.7) such that $H_{i^{**}}(\bar{q}, y') < 0$ for the node $i^{**} \in \Pi$ constructed above. Since the objective function of (3.7) does not depend on y but just on q, it follows from the fact that that $(\bar{q}, \bar{y})$ was an optimal solution to (3.7), that $(\bar{q}, y')$ is also an optimal solution to (3.7). But now, using the already proven fact that demand satisfaction comes as an equality at solutions to (3.7) for all active generators, we infer from the relation $i^{**} \leq l$ shown above that $H_{i^{**}}(\bar{q}, y') = 0$, a contradiction. Consequently, our original assumption $H_{i^*}(\bar{q}, \bar{y}) < 0$ is wrong showing that $H_i(\bar{q}, \bar{y}) = 0$ for $i \in \{l+1, \ldots, N\}$. This completes the proof of the lemma. □

The surjectivity condition in statement 1. (c) of Lemma 7.2 can be interpreted as follows in a special case: if $l = N - 1$ (all generators but one are active), then there must exist at least one non-congested transmission line leading to the non-active generator.

Proposition 7.1 (**CRCQ in lossless case**). *Let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution to (3.9) satisfying (7.1). If $\rho_j = 0$ $(j = 1, \ldots, m)$, then CRCQ holds at $(\bar{q}, \bar{y}) \in G$.*

Proof. Since LICQ implies CRCQ, we only need to investigate cases where LICQ does not hold. Let $1 \leq l < N$ and $0 \leq k < m$ and recall $\nabla H(\bar{q}, \bar{y})$ as provided in (7.4) with $\rho_j = 0$

$$\nabla H(\bar{q}, \bar{y}) = \begin{pmatrix} -I_1 & 0 & -B^{11} & -B^{12} \\ 0 & -I_2 & -B^{21} & -B^{22} \\ 0 & -I_2 & 0 & 0 \\ 0 & 0 & 0 & I_3 \end{pmatrix}.$$

Then since $\nabla H(\bar{q}, \bar{y})$ neither depends on $\bar{q}$ nor $\bar{y}$, there will always exist a neighborhood $\mathcal{U}$ of $(\bar{q}, \bar{y})$ such that $\operatorname{rank}\{\nabla H_I(q, y)\}$ remains constant for all $(q, y) \in \mathcal{U}$ and any $I \subseteq \{1, \ldots, 2N + m - l - k\}$. □

Unfortunately, obtaining a CRCQ-result for the EPEC where losses are accounted for can only be done on a case-by-case basic and by restricting ρ and $\hat{y}$.

Next, we identify situations in which solutions of the generalized equation arising from the first order optimality conditions of the ISO problem satisfy SSOSC at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$. We recall that SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$, if

$$\left\langle d, \nabla_{(q,y)} \mathcal{L}(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}, \bar{\xi}) d \right\rangle > 0 \quad \forall \xi, \forall d \neq 0 : d \in \ker\{\nabla H_{I_+(\bar{q}, \bar{y}, \bar{\xi})}(\bar{q}, \bar{y})\}. \quad (7.18)$$

Here,

$$\mathcal{L}(\alpha, \beta, q, y, \xi) := \begin{pmatrix} \alpha + 2[\operatorname{diag} \beta] q \\ 0 \end{pmatrix} + \nabla^T H(q, y) \xi$$

i.e., the vector Lagrangian associated with (3.8) and $\bar{\xi}$ denotes any Lagrange multiplier associated with the constraint mapping H at the solution $(\bar{q}, \bar{y})$.

Proposition 7.2 (**SSOSC for the ISO problem**). *Let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution to (3.9) satisfying (7.1). Then there exists some $\Delta > 0$ such that under the condition $\rho_j \in (0, \Delta)$ $(j = 1, \ldots, m)$, the following holds true: if $l = N$ (all generators active) or $k = m$ (no congestion), then*

1. *SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$, i.e., (7.18) is satisfied.*

Moreover, the same conclusion can also be drawn if $\rho_j = 0$ $(j = 1, \ldots, m)$ under the additional assumption that the network graph is a tree (i.e., it does not contain cycles).

Proof. Exploiting the explicit structure of $\mathcal{L}$ and H, one calculates

$$\nabla_{(q,y)}\mathcal{L}(\bar{\alpha},\bar{\beta},\bar{q},\bar{y},\bar{\xi}) = \begin{pmatrix} 2[\operatorname{diag}\bar{\beta}] & 0 & \cdots & 0 \\ 0 & \rho_1 \sum_{i=1}^{N} \bar{\xi}_i |b_{i1}| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \rho_m \sum_{i=1}^{N} \bar{\xi}_i |b_{im}| \end{pmatrix}.$$

This diagonal matrix contains only positive entries. Indeed, this is clear for the first block because $\bar{\beta}_i > 0$ (see (7.1)). For the remaining entries

$$\rho_j \sum_{i=1}^{N} \bar{\xi}_i |b_{ij}| \tag{7.19}$$

note that $\rho_j > 0$ by assumption. Moreover, $\bar{\xi}_i > 0$ for $i = 1,\ldots,N$. Indeed, recall that $\left(\bar{\xi}_1,\ldots,\bar{\xi}_l\right)$ and $\left(\bar{\xi}_{l+1},\ldots,\bar{\xi}_N\right)$ correspond to the Lagrange multipliers of the demand satisfaction relations (first N components of H) for active and non-active generators, respectively. Then statements 3. (a) and (b) of Lemma 7.2 indicate that strict complementarity (i.e., $\bar{\xi}_i > 0$ for $i = 1,\ldots,N$) holds whenever $l = N$ or $k = m$. On the other hand, as our network is a connected graph, for each index j, there exists at least one (exactly: two) i such that $|b_{ij}| = 1$. Consequently, (7.19) is strictly positive for all $j = 1,\ldots,m$. Evidently, $\nabla_{(q,y)}\mathcal{L}(\bar{\alpha},\bar{\beta},\bar{q},\bar{y},\bar{\xi})$ is positive definite and (7.18) is satisfied.

For the second assertion, assume that $\rho_j = 0$ $(j = 1,\ldots,m)$ and the network does not contain any cycles. Then,

$$\nabla_{(q,y)}\mathcal{L}(\bar{\alpha},\bar{\beta},\bar{q},\bar{y},\bar{\xi}) = \begin{pmatrix} 2[\operatorname{diag}\bar{\beta}] & 0 \\ 0 & 0 \end{pmatrix}.$$

Choose an arbitrary d as indicated in (7.18). In particular, $d \neq 0$. Moreover, as already shown above, one has that $\bar{\xi}_i > 0$ for $i = 1,\ldots,N$ holds whenever $l = N$ or $k = m$. Consequently, by (7.18), $\nabla H_i(\bar{q},\bar{y})d = 0$ for $i = 1,\ldots,N$. Using the partition $d = (d_1, d_2)$, the concrete shape of $\nabla H(\bar{q},\bar{y})$ yields that $d_1 + Bd_2 = 0$. However, since B is injective as the incidence matrix of a tree, $d \neq 0$ already implies that $d_1 \neq 0$. But then,

$$\left\langle d, \nabla_{(q,y)}\mathcal{L}(\bar{\alpha},\bar{\beta},\bar{q},\bar{y},\bar{\xi})d\right\rangle = 2\left\langle d_1, [\operatorname{diag}\bar{\beta}]d_1\right\rangle > 0$$

as was to be shown in (7.18). □

7.2 A Remark on Existence of Solutions

As a consequence of Proposition 7.2 and Lemma 7.2, the solutions to the ISO problem (3.7) can often be parameterized by a single-valued and Lipschitzian mapping $(q(\alpha,\beta), y(\alpha,\beta))$, i.e., the equilibrium constraint is strongly regular at $(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$. In such cases, this allows us to locally replace the EPEC (3.9) by a classical Nash-Cournot game, thus making it amenable to what is called the

Implicit Programming Approach (see e.g., Luo et al. [1997, Chapter 4 Section 2] and Outrata et al. [1998, Chapter 7])

$$\min_{(\alpha_i,\beta_i)\in\mathbb{R}^2} -f_i(\alpha_i, \beta_i, q\,(\alpha, \beta)\,, y\,(\alpha, \beta)) \quad (i = 1, \ldots, N).$$

Nevertheless, this approach is not possible in all relevant cases. For instance, if the network graph contains cycles, as it is typically the case, then Proposition 7.2 does not apply to the loss-free model ($\rho_j = 0$ $(j = 1, \ldots, m)$).

Given that strong regularity implies the existence of neighborhoods $\mathcal{U}$ of $(\bar{\alpha}, \bar{\beta})$ and $\mathcal{V}$ of $(\bar{q}, \bar{y})$ on which $q(\alpha, \beta)$ and $y(\alpha, \beta)$ are locally Lipschitz, it seems reasonable that this (local) reformulation of (3.9) into a traditional Nash equilibrium problem might lead to an argument for proving the existence of a local EPEC solution. The main difficulty in doing so arises from the fact that each f_i is most likely nothing more than locally Lipschitz. Thus, the classical existence results of Nash [1951], Debreu [1952], Fan [1984] et cetera, do not apply. However, it may be possible to apply results pertaining to non-quasiconvex objective functions such as are provided in Baye et al. [1993] and Nishimura and Friedman [1981]. Even if these results cannot help in proving the existence of an EPEC solution in the spot market model, it appears as though this may be one approach which one could take with other classes of EPECs where the objective functions enjoy better properties.

7.3 Verifying Calmness

In order to derive M-stationarity conditions for the spot market EPEC, we need to verify that the perturbation mappings

$$\Psi_i(u) := \left\{(\alpha_i, \beta_i, q, y) \,\middle|\, u \in F(\bar{\alpha}_{-i}, \alpha_i, \bar{\beta}_{-i}, \beta_i, q, y) + N_G(q, y)\right\} \tag{7.20}$$

are calm at $\left(0, \bar{\alpha}_i, \bar{\beta}_i, \bar{q}, \bar{y}\right)$.

According to a well-known result by Robinson [1981, Proposition 1], a multifunction with a polyhedral graph (i.e., the graph of which is a finite union of polyhedra) is calm at any point of its graph. Hence, the simplest way to verify calmness of (7.20) consists in checking the polyhedrality of its graph. Evidently, if the mapping F is linear and transmission losses are ignored, then the mappings Ψ_i are calm at all points of their graphs. Indeed, in the loss-free case the feasible set G of the ISO problem (3.7) becomes a polyhedron, thus making the graph of the mapping N_G a finite union of polyhedra. Given the linearity of F, the graph of Ψ_i is also a finite union of polyhedra. Unfortunately, the mappings

$$F(\bar{\alpha}_{-i}, \alpha_i, \bar{\beta}_{-i}, \beta_i, q, y) = \begin{pmatrix} (\bar{\alpha}_{-i}, \alpha_i) + 2[\mathrm{diag}\,\left(\bar{\beta}_{-i}, \beta_i\right)]q \\ 0 \end{pmatrix}$$

are not linear in our case because of the bilinear term $\beta_i q_i$. Things would be different under the special assumption of *partial bidding* made in Hu and Ralph

[2005], Hu et al. [2007]: here, the producers quadratic cost term is assumed to be known by every market participant and thus, it is not a part of the decision variables. In such a case, F becomes linear and calmness of (7.20) can be taken for granted in the loss-free case (this fact was exploited, for instance in Henrion and Römisch [2007]). However, we do not wish to make such a restrictive assumption in this thesis.

As observed in Chapter 5, when the multifunctions in question are nonpolyhedral, another way of verifying calmness is to check if the stronger Aubin property holds by using available criteria. Indeed, we have the following result:

Proposition 7.3 (**calmness via strong regularity**). *Under the assumptions of Proposition 7.2, the perturbation mappings Ψ_i in (7.20) have the Aubin property and, hence, are calm at $\left(0, \bar{\alpha}_i, \bar{\beta}_i, \bar{q}, \bar{y}\right)$.*

Proof. Proposition 7.2 implies that SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ in the considered cases, i.e., either $l = N$, $k = m$, or $\rho_j = 0$ $(j = 1, \ldots, m)$ and B induces a tree. Moreover, it is easy to see that in each of these cases, statement 1. (b) of Lemma 7.2 holds, i.e., $\nabla H(\bar{q}, \bar{y})$ is surjective. Referring to the discussion at the end of Chapter 4, if SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ for the generalized equation

$$0 \in F(\alpha, \beta, q, y) + N_G(q, y),$$

then it does as well for the more restricted generalized equation

$$0 \in F(\bar{\alpha}_{-i}, \alpha_i, \bar{\beta}_{-i}, \beta_i, q, y) + N_G(q, y)$$

at $(\bar{\alpha}_i, \bar{\beta}_i, \bar{q}, \bar{y})$. Then by Corollary 5.1, the Ψ_i as defined in (7.20) have the Aubin property at $(0, \bar{\alpha}_i, \bar{\beta}_i, \bar{q}, \bar{y})$ for all $i = 1, \ldots, N$ and are therefore calm there as well. □

Note that Proposition 7.3 cannot be applied in the loss-free case whenever the network graph contains cycles, as the conditions in Proposition 7.2 are violated. For illustration we provide the following example demonstrating that in such cases the perturbation mappings Ψ_i in (7.20) do not have the Aubin property.

Example 7.1 (**failure of the Aubin property**). *Given the spot market EPEC (3.9), let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution such that (7.1) is satisfied. Furthermore, let $N = 3$, $m = 3$, $l = 3$, $k = 3$, $\rho_j = 0$ $(j = 1, 2, 3)$, i.e., there is neither congestion nor transmission losses nor non-active generators, and define a cyclic graph induced via the incidence matrix*

$$B = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

Since $l = N$, Lemma 7.2 indicates that $\nabla H(\bar{q}, \bar{y})$ is surjective. Then we can rewrite the normal cone $N_G(\bar{q}, \bar{y})$ (cf. Rockafellar and Wets [1998, Theorem 6.14])

$$N_G(\bar{q}, \bar{y}) = \nabla^T H(\bar{q}, \bar{y}) N_{\mathbb{R}^3_-}(H(\bar{q}, \bar{y})).$$

From $\nabla H(\bar{q},\bar{y}) = (-I| - B)$ *(see (7.4) with the special data of this example) and from the concrete shapes of* F *and* B *we derive that, for any* $i \in \{1,\dots,N\}$,

$$u = (u_1,\dots,u_6) \in F(\bar{\alpha}_{-i},\alpha_i,\bar{\beta}_{-i},\beta_i,q,y) + N_G(\bar{q},\bar{y}) \Longrightarrow$$

$$\exists\bar{\lambda} = \left(\bar{\lambda}_1,\bar{\lambda}_2,\bar{\lambda}_3\right) \in N_{\mathbb{R}^3_-}(H(\bar{q},\bar{y})) : \begin{array}{l} u_4 = \bar{\lambda}_1 - \bar{\lambda}_2 \\ u_5 = \bar{\lambda}_2 - \bar{\lambda}_3 \\ u_6 = \bar{\lambda}_3 - \bar{\lambda}_1. \end{array}$$

Consequently, if $u \in F(\bar{\alpha}_{-i},\alpha_i,\bar{\beta}_{-i},\beta_i,q,y) + N_G(\bar{q},\bar{y})$, *then necessarily* $u_4+u_5+u_6 = 0$. *By contraposition, if* $u_4+u_5+u_6 \neq 0$ *for some* u, *then necessarily* $\Psi_i(u) = \emptyset$ *for the multifunctions defined in (7.20). As one may now easily construct a sequence* $u^{(n)} \to 0$ *with* $\Psi_i(u^{(n)}) = \emptyset$, *it follows that* Ψ_i *cannot have the Aubin property at* $\left(0,\bar{\alpha}_i,\bar{\beta}_i,\bar{q},\bar{y}\right)$.

It follows that in the loss-free case one can neither rely on a polyhedrality argument nor on a verification of the Aubin property in order to verify calmness. Fortunately, the loss of polyhedrality turns out to be weak enough to allow a direct verification of the calmness of Ψ_i. That is, Theorem 5.1 can be applied, which leads to the next proposition.

Proposition 7.4 (**direct verification of calmness via Theorem 5.1**). *Let* $(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$ *be a solution to (3.9) satisfying (7.1). If for all* $j = 1,\dots,m$, $\rho_j = 0$, *then for all* $i = 1,\dots,N$ *the perturbation mappings* Ψ_i *defined in (7.20) are calm at* $(0,\bar{\alpha}_i,\bar{\beta}_i,\bar{q},\bar{y})$.

Proof. Being that Ψ_i arises from the first order optimality conditions of the ISO problem, we can use Theorem 5.1. Without loss of generality, let $i = 1$, then

$$F(\bar{\alpha}_{-1},\alpha_1,\bar{\beta}_{-1},\beta_1,q,y) = \begin{pmatrix} \alpha_1 + 2\beta_1 q_1 \\ \bar{\alpha}_{-1} + 2[\operatorname{diag}\bar{\beta}_{-1}]q_{-1} \\ 0 \end{pmatrix}.$$

Thus, $F(\bar{\alpha}_{-1},\alpha_1,\bar{\beta}_{-1},\beta_1,q,y)$ has the form:

$$F(\bar{\alpha}_{-1},\alpha_1,\bar{\beta}_{-1},\beta_1,q,y) = \begin{pmatrix} \Delta_1(\alpha_1,\beta_1,q,y) \\ \Delta_2(q,y) \end{pmatrix},$$

where $\Delta_1(\alpha_1,\beta_1,q,y) = \alpha_1 + 2\beta_1 q_1$ and

$$\Delta_2(q,y) = \begin{pmatrix} \bar{\alpha}_{-1} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 2[\operatorname{diag}\bar{\beta}_{-1}] & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_{-1} \\ y \end{pmatrix}.$$

Then assumption 1. of Theorem 5.1 is fulfilled if we substitute $(\alpha_1,\beta_1) = u_1$, $u = u_2$, $(q,y) = z$, $t_1 = 1$, $t_2 = N-1+m$, and $C = G$. Moreover, G is polyhedral and convex, as $H(q,y)$ is affine linear due to the fact that $\rho_j = 0$ for all $j = 1,\dots,m$. Then Lemma 7.2 (statement 1. (a)) ensures the existence of a

Slater point $(\tilde{q}, \tilde{y}) \in G$. Thus, assumption 2. is satisfied. Finally, by noting that

$$\nabla_{(\alpha_1,\beta_1)}\Delta_1(\bar{\alpha}_1, \bar{\beta}_1, \bar{q}, \bar{y}) = (1, 2\bar{q}_1) \neq (0, 0),$$

we see that assumption 3. is fulfilled. Therefore, Theorem 5.1 implies Ψ_1 is calm at $(0, \bar{\alpha}_1, \bar{\beta}_1, \bar{q}, \bar{y})$. $\square$

Summarizing, we have the following tables, which compile the regularity and stability results demonstrated in this chapter.

$\rho > 0$	Constraint Qualification	Stability of Ψ_i
$N = l$, $m = k$	LICQ, SSOSC	Aubin Property
$N = l$, $m > k$	LICQ, SSOSC	Aubin Property
$N > l$, $m = k$	LICQ, SSOSC	Aubin Property
$N > l$, $m > k$	MFCQ/LICQ†	-

We use '$\dagger$' to remind the reader that this only holds when B^{21} is surjective (see Lemma 7.2 statement 1(c)).

$\rho = 0$	Constraint Qualification	Stability of Ψ_i
$N = l$, $m = k$	LICQ, SSOSC*	Calm/Aubin Prop.*
$N = l$, $m > k$	LICQ, SSOSC*	Calm/Aubin Prop.*
$N > l$, $m = k$	LICQ, SSOSC*	Calm/Aubin Prop.*
$N > l$, $m > k$	MFCQ,CRCQ/LICQ†, SSOSC*	Calm/Aubin Prop.†,*

Here, '*' is to remind the reader that SSOSC and, by way of Corollary 5.1, the Aubin property for Ψ_i only hold when B generates an acyclic graph.

Chapter 8

Explicit Characterizations of Solutions using Dual Stationarity Conditions

At this point in the thesis, we have developed all the tools and results needed in order to derive explicit stationarity conditions and to use these conditions to characterize solutions to the spot market EPEC (3.9). Though part of this chapter can be thought of as a comparison of the selectivity of the stationarity conditions outlined in Chapter 4, its main purpose is to demonstrate how the results from the previous chapters can be applied in a practical setting. We end this chapter with a section containing an analysis of the stochastic spot market EPEC, e.g., its structural properties and explicit stationarity conditions. This is done by considering a discrete probability measure, which allows us to reformulate the SEPEC into a larger, yet deterministic, EPEC.

8.1 Explicit M-stationarity Conditions for the Spot Market EPEC

We begin with the following theorem in the context of the spot market EPEC (see Outrata [2004, Theorem 3.1]).

Theorem 8.1 (M-stationarity for the spot market EPEC). *Let $\left(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}\right)$ be a solution to (3.9). If for all $i = 1, \ldots, N$, the multifunctions*

$$\Psi_i(u) := \left\{(\alpha_i, \beta_i, q, y) \,\middle|\, u \in F(\bar{\alpha}_{-i}, \alpha_i, \bar{\beta}_{-i}, \beta_i, q, y) + N_G(q, y)\right\}$$

are calm at $\left(0, \bar{\alpha}_i, \bar{\beta}_i, \bar{q}, \bar{y}\right)$, then for all $i = 1, \ldots, N$, there exist v^i such that the following relations hold

$$0 = \nabla_{\alpha_i,\beta_i} f_i\left(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}\right) + \nabla^T_{\alpha_i,\beta_i} F\left(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}\right) v^i \tag{8.1}$$

$$0 \in \nabla_{q,y} f_i\left(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}\right) + \nabla^T_{q,y} F\left(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}\right) v^i + D^* N_G(\bar{q}, \bar{y}, \bar{z})(v^i) \tag{8.2}$$

where $\bar{z} = -F\left(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}\right)$

As noted at the end of Chapter 4, under the right conditions, we can derive more explicit versions of (8.1) and (8.2) by following an argument similar to that which was used in Corollary 4.1 for each $i = 1, \dots, N$. We demonstrate this in the next result.

Proposition 8.1 (Explicit M-stationarity Conditions). *Let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution to (3.9) satisfying (7.1). Assume that $l = N$ (all generators active) or $k = m$ (no congestion). Then there exists some $\Delta > 0$ such that under the condition $\rho_j \in (0, \Delta)$ $(j = 1, \dots, m)$ (small positive losses) or $\rho_j = 0$ $(j = 1, \dots, m)$ (no losses) the following holds true: there exists a unique $\bar{\lambda} \in \mathbb{R}_+^{2N+m-l-k}$ and for all $i = 1, \dots, N$, there exist $(v^i, w^i) \in \mathbb{R}^{N+m} \times \mathbb{R}^{2N+m-l-k}$ such that*

$$\nabla_{\alpha_i,\beta_i} f_i(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}) = \nabla^T_{\alpha_i,\beta_i} F(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}) v^i \tag{8.3}$$

$$\nabla_{q,y} f_i(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}) = \nabla^T_{q,y} F(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}) v^i + \left(\sum_{j=1}^{N} \bar{\lambda}_j \nabla^2 H_j(\bar{q}, \bar{y}) \right) v^i + \nabla^T H(\bar{q}, \bar{y}) w^i \tag{8.4}$$

$$\nabla H_j(\bar{q}, \bar{y}) v^i = 0 \quad \forall j : \bar{\lambda}_j > 0 \tag{8.5}$$

$$w^i = 0 \quad \forall j : \bar{\lambda}_j = 0,\ \nabla H_j(\bar{q}, \bar{y}) v^i < 0 \tag{8.6}$$

$$w^i \geq 0 \quad \forall j : \bar{\lambda}_j = 0,\ \nabla H_j(\bar{q}, \bar{y}) v^i > 0 \tag{8.7}$$

$$F(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y}) = -\nabla^T H(\bar{q}, \bar{y}) \bar{\lambda}. \tag{8.8}$$

Proof. Note first that in all indicated constellations, $\nabla H(\bar{q}, \bar{y})$ is surjective according to Lemma 7.2, statement 1.(b) and that the calmness assumption for the multifunctions (7.20) is satisfied by virtue of Propositions 7.3 (for the case $\rho_j \in (0, \Delta)$) and 7.4 (for the case $\rho_j = 0$). Thus for each $i = 1, \dots, N$, we may apply Theorem 8.1 in conjunction with Corollary 4.1, which upon setting $s = 2N$,$t = N+m$, $p = 2N+m-l+k$, $x = (\alpha_i, \beta_i)$, $z = (q, y)$, $F(x, z) = F(\bar{\alpha}_{-i}, \alpha, \bar{\beta}_{-i}, \beta_i, q, y)$ and $A(z) = H(q, y)$ yields the existence of $(v^i, w^i) \in \mathbb{R}^{N+m} \times \mathbb{R}^{2N+m-l+k}$ such that (8.3)-(8.8) hold. □

Given Proposition 8.1, it is easy to see how we can carry over many of the results pertaining to stationarity conditions found in Chapter 4 to the EPEC setting. Also note that in the case where the vectors $\nabla H_j(\bar{q}, \bar{y})$ are only positive linearly independent, the calmness result in Proposition 7.4 is still applicable, as long as we are considering the lossless EPEC. Furthermore, in the lossless case, Proposition 7.1 indicates that CRCQ holds at $(\bar{q}, \bar{y})$. Therefore, in such a setting, we could again derive a set of explicit M-stationarity conditions using Theorem 8.1 in conjunction with Corollary 4.2. Finally, in the event that B generates an acyclic network, Proposition 7.2 indicates that SSOSC holds, in which case we could derive CM-stationarity conditions, though the calculation of $D^* N_K$ would be required.

We now provide an even more explicit set of stationarity conditions than (8.3)-(8.8) for a select type of settings to be used afterwards in the examples section. We know that in the absence of losses and congestion, the Jacobian of the mapping H takes the form (see (7.4))

$$\nabla H(\bar{q},\bar{y}) = \begin{pmatrix} -I_1 & 0 & -B^1 \\ 0 & -I_2 & -B^2 \\ 0 & -I_2 & 0 \end{pmatrix}. \tag{8.9}$$

Proposition 8.2. *Let $(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$ be a solution to (3.9) satisfying (7.1). Assume that $\rho_j = 0$ for $j = 1,\ldots,m$ (no losses) and $k = m$ (no congestion). Then there exist a unique $\bar{\lambda} \in \mathbb{R}_+^{2N-l}$ and for all $i = 1,\ldots,N$, there exist $(v^i,w^i) \in \mathbb{R}^{N+m} \times \mathbb{R}^{2N-l}$ such that*

$$(v^i)_i = \begin{cases} \bar{q}_i & \text{if } i \in \{1,\ldots,l\} \\ 0 & \text{if } i \in \{l+1,\ldots,N\} \end{cases} \tag{8.10}$$

$$v_a^i + B^1 v_c^i = v_b^i + B^2 v_c^i = 0 \tag{8.11}$$

$$(v^i)_{l+j} = 0 \quad \forall j \in \{1,\ldots,N-l\} : \bar{\lambda}_{N+j} > 0 \tag{8.12}$$

$$(w^i)_i = \gamma_i - \bar{\alpha}_i + 2(\delta_i - \bar{\beta}_i)\bar{q}_i \quad \text{if } i \in \{1,\ldots,l\} \tag{8.13}$$

$$(w^i)_i + (w^i)_{N+i-l} = \gamma_i - \bar{\alpha}_i \quad \text{if } i \in \{l+1,\ldots,N\} \tag{8.14}$$

$$(w^i)_j = 2\bar{\beta}_j (v^i)_j \quad \text{if } j \in \{1,\ldots,l\},\ j \neq i \tag{8.15}$$

$$(w^i)_j + (w^i)_{N+j-l} = 2\bar{\beta}_j (v^i)_j \quad \text{if } j \in \{l+1,\ldots,N\},\ j \neq i \tag{8.16}$$

$$(w^i)_{N+j-l} = 0 \quad \text{if } j \in \{l+1,\ldots,N\},\ \bar{\lambda}_{N+j-l} = 0,\ (v^i)_j > 0 \tag{8.17}$$

$$(w^i)_{N+j-l} \geq 0 \quad \text{if } j \in \{l+1,\ldots,N\},\ \bar{\lambda}_{N+j-l} = 0,\ (v^i)_j < 0 \tag{8.18}$$

$$(B^1)^T w_a^i + (B^2)^T w_b^i = 0 \tag{8.19}$$

$$\bar{\alpha}_1 + 2\bar{\beta}_1\bar{q}_1 = \bar{\alpha}_j + 2\bar{\beta}_j\bar{q}_j \quad (j = 1,\ldots,l) \tag{8.20}$$

$$= \bar{\alpha}_{j'} - \bar{\lambda}_{N+j'-l} \quad (j' = l+1,\ldots,N) \tag{8.21}$$

Here $(\cdot)_j$ identifies a concrete component of a vector, whereas lower indices 'a', 'b', 'c' obey the partition of the Jacobian in (8.9) and its transpose, respectively.

Proof. (8.10) follows from (8.3) upon calculating $\nabla_{\alpha_i,\beta_i} f_i(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$ and $\nabla^T_{\alpha_i,\beta_i} F(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$ and recalling that $\bar{q}_i > 0$ for $i \leq l$ and $\bar{q}_i = 0$ for $i \in \{l+1,\ldots,N\}$. Next, observe that due to our assumption $k = m$, we know that, by statement 3.(b) of Lemma 7.2, $\bar{\lambda}_j > 0$ for $j = 1,\ldots,N$. Therefore, (8.11) and (8.12) are implied by (8.5) taking into account the shape of the Jacobian in (8.9). Relations (8.13)-(8.16) and (8.19) are derived from (8.4) (with the Hessian term missing due to linearity as a consequence of the loss-free case) upon calculating $\nabla_{q,y} f_i(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$ and $\nabla^T_{q,y} F(\bar{\alpha},\bar{\beta},\bar{q},\bar{y})$ and comparing components in the light of the shape of the transposed Jacobian in (8.9). Relations (8.17) and (8.18) correspond to (8.6) and (8.7), where again $\bar{\lambda}_j > 0$ for $j = 1,\ldots,N$ and the shape of the third row block in (8.9) are exploited. Relations (8.20) and (8.21) result from the KKT- conditions of the ISO-problem (3.7) along the lines of relations (7.8), (7.9), (7.10) and (7.11). Indeed, (7.10) and (7.11) imply, due to the absence of losses, that

$$(\bar{\lambda}_1, \ldots, \bar{\lambda}_N)^T \in \ker B^T$$

which in turn results in $\bar{\lambda}_1 = \cdots = \bar{\lambda}_N$ by virtue of Lemma 7.1, statement 1. Now, (8.20) and (8.21) follow from (7.8) and (7.9) with an appropriate change of notation. □

8.2 Examples with Two Settlements

In the following, we illustrate how to use (8.3)-(8.8), and, where applicable, (8.10)-(8.21) for analyzing various types of solutions to the smallest relevant spot market settings.

Example 8.1 (M-stationarity in a Nonlinear/Nonsmooth Setting). Suppose $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ is a solution to (3.9) satisfying (7.1) such that

$$N = l = 1, \quad m = k = 1, \quad \rho > 0, \quad B = (-1, 1)^T, \quad -1 < \rho\bar{y} < 1$$

that is, $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ is a solution such that only the first producer is active, the line is uncongested, transmission losses are considered, and power is flowing from node 1 to node 2. Such a solution fulfills all the necessary criteria in order to apply Theorem 8.1. In this setting, it is easy to show that (8.3)-(8.8) yield no information for $i = 2$, however, for $i = 1$ they become:

$$(v^1)_1 = \bar{q}_1 > 0 \tag{8.22}$$

$$(\bar{\alpha}_1 - \gamma_1) + 2(\bar{\beta}_1 - \delta_1)\bar{q}_1 = -(w^1)_1 \tag{8.23}$$

$$2\bar{\beta}_2(v^1)_2 = (w^1)_2 + (w^1)_3, \tag{8.24}$$

$$0 = (1 + \rho\bar{y})(w^1)_1 + (\rho\bar{y} - 1)(w^1)_2, \tag{8.25}$$

$$(v^1)_1 = (1 + \rho\bar{y})(v^1)_3 \tag{8.26}$$

$$(v^1)_2 = (\rho\bar{y} - 1)(v^1)_3 \tag{8.27}$$

$$0 < \bar{\alpha}_1 + 2\bar{\beta}_1\bar{q}_1 = \bar{\lambda}_1 \tag{8.28}$$

$$0 < \bar{\alpha}_2 = \bar{\lambda}_2 - \bar{\lambda}_3 \tag{8.29}$$

$$0 = (1 + \rho\bar{y})\bar{\lambda}_1 + (\rho\bar{y} - 1)\bar{\lambda}_2 \tag{8.30}$$

Note first that (8.26) and (8.27) result from (8.5), since Lemma 7.2 statement 3(a) implies $I_+ \supseteq \{1, 2\}$. We claim this holds as an equality. Indeed, if we assume $\bar{\lambda}_3 > 0$, then by (8.5), it must hold that $\nabla H_3(\bar{q}, \bar{y})v^1 = 0$, which upon referring to the structure of $\nabla H(\bar{q}, \bar{y})$ (see (7.4)) implies $(v^1)_2 = 0$. However, due to the condition on ρ and $\bar{y}$, this implies via (8.26) and (8.27) that $(v^1)_1 = 0$, a contradiction. Thus, $I_+ = \{1, 2\}$ and $I_0 = \{3\}$.

Using now (8.22),(8.26), and (8.27), we observe that $(v^1)_2 = \kappa(\rho, \bar{y})\bar{q}_1 < 0$, where $\kappa(\rho, \bar{y}) := (\rho\bar{y} - 1)/(1 + \rho\bar{y})$. Then $\nabla H_{I_0}(\bar{q}, \bar{y})v^1 = \nabla H_3(\bar{q}, \bar{y})v^1 > 0$.

Hence, $(w^1)_3 \geq 0$ (see (8.7)). Note that when $i = 2$, we can proceed analogously, resulting in $\nabla H(\bar{q}, \bar{y})v^2 = 0$. Thus Theorem 4.3 cannot be applied for obtaining S-stationarity conditions[1].

Finally, by Lemma 7.2 statement 2, we know $H(\bar{q}, \bar{y}) = 0$. Then we can derive the following relations

$$
\begin{aligned}
0 &\leq (\bar{\alpha}_1 - \gamma_1) + 2(\bar{\beta}_1 - \kappa(\rho, \bar{y})^2 \bar{\beta}_2 - \delta_1)\bar{q}_1 \\
\bar{\alpha}_2 &= \frac{\bar{\alpha}_1 + 2\bar{\beta}_1 \bar{q}_1}{-\kappa(\rho, \bar{y})} \\
\bar{q}_1 &= \frac{-\kappa(\rho, \bar{y})\bar{\alpha}_2 - \bar{\alpha}_1}{2\bar{\beta}_1} = d + \rho \bar{y}^2
\end{aligned}
$$

These relations thus classify all solutions of the type considered. Furthermore, any feasible solution to the above three relations is an M-stationary point.

Example 8.1 shows how the conditions (8.3)-(8.8) can sometimes be reduced to a collection of inequalities and equalities without extra multipliers. In the sequel, we will take advantage of this fact for our analysis.

Example 8.1 (Strengths and Weaknesses of M-stationarity). *Consider again a solution of the type used in Example 8.1, however for this example, we assume $\rho = 0$. In this case, we can use Proposition 8.2, in which case (8.10)-(8.21) reduce to the following for $i = 1$.*

$$
\begin{aligned}
(v^1)_1 &= \bar{q}_1 > 0 && (8.31) \\
(\bar{\alpha}_1 - \gamma_1) + 2(\bar{\beta}_1 - \delta_1)\bar{q}_1 &= -(w^1)_1 && (8.32) \\
2\bar{\beta}_2 (v^1)_2 &= (w^1)_2 + (w^1)_3, && (8.33) \\
0 &= (w^1)_1 - (w^1)_2, && (8.34) \\
(v^1)_1 &= (v^1)_3 && (8.35) \\
(v^1)_2 &= -(v^1)_3 && (8.36) \\
0 < \bar{\alpha}_1 + 2\bar{\beta}_1 \bar{q}_1 &= \bar{\lambda}_1 && (8.37) \\
0 < \bar{\alpha}_2 &= \bar{\lambda}_2 - \bar{\lambda}_3 && (8.38) \\
0 &= \bar{\lambda}_1 - \bar{\lambda}_2 && (8.39)
\end{aligned}
$$

As in the previous example, it is easy to argue that $I_+ = \{1, 2\}$ and $I_0 = \{3\}$, which again leads to $\nabla H_3(\bar{q}, \bar{y})v^1 > 0$ (and $\nabla H_3(\bar{q}, \bar{y})v^2 = 0$). Then (8.31)-(8.39) reduce to

$$
\begin{aligned}
(\bar{\alpha}_1 - \gamma_1) + 2(\bar{\beta}_1 - \bar{\beta}_2 - \delta_1)\bar{q}_1 &\geq 0 && (8.40) \\
\bar{\alpha}_1 + 2\bar{\beta}_1 \bar{q}_1 &= \bar{\alpha}_2 && (8.41)
\end{aligned}
$$

[1] However, this still indicates that Theorem 4.3 can be applied in order to argue that the M-stationarity conditions for the MPEC corresponding to generator 1 are in fact S-stationarity conditions.

Using Lemma 7.2 statement 2, i.e., $H(\bar{q}, \bar{y}) = 0$*, we see that* $\bar{q}_1 = d$*. So in such a situation, generator 1 is producing enough electricity to cover the demand in both nodes of the network. This is similar to the previous example, only that there, the amount of electricity lost due to transmission had to be taken into account as well. Now, it is easy to derive the relation:*

$$\bar{\alpha}_2 \geq \gamma_1 + 2(\bar{\beta}_2 - \delta_1)d. \tag{8.42}$$

Given $\bar{\beta}_2, d > 0$*, we see that*

$$\bar{\alpha}_2 > \gamma_1 - 2\delta_1 d,$$

which implies that if the linear bid coefficient of generator 2 is strictly greater that $\gamma_1 - 2\delta_1 d$*, then it is possible for generator 2 to be forced from the market. Of course, if the second player were to bid* $\bar{\alpha}_2 \leq \gamma_1 - 2\delta_1 d$*, then no such solution would be possible as the M-stationarity conditions would be violated. Not only does this demonstrate how the stationarity conditions can be also used to* rule out *possible solutions but also, it provides a quantitative statement regarding the decision variables and parameters of the model.*

Example 8.2 (Example 8.1 continued: selectivity of M-stationarity). *We now examine the selectivity of M-stationarity by using relations (8.40) and (8.41). Suppose that*

$$\gamma_1 = 1, \quad \delta_1 = 0.25, \quad \gamma_2 = 2, \quad \delta_2 = 1,$$

then it is easy to see that

$$\bar{\alpha}_1 = 1, \quad \bar{\beta}_1 = 0.5, \quad \bar{\alpha}_2 = 2, \quad \bar{\beta}_2 = 0.25$$

satisfy (8.40) and (8.41). In fact, by plotting the profit functions at this point (see Figure 8.1), we see that we have identified a solution and not just an M-stationary point.

Figure 8.1: Profit functions at $(\bar{\alpha}_1, \bar{\beta}_1, \bar{\alpha}_2, \bar{\beta}_2) = (1, 0.5, 2, 0.25)$

Here, generator 1's profit function is on the left and generator 2's profit function is on the right. Notice how the maxima occur at points in which the profit functions are nonsmooth. *This provides a concrete illustration of the non-differentiable nature of EPECs, despite the fact that we have smooth data. Consider now that*

$$\bar{\alpha}_1 = 1, \quad \bar{\beta}_1 = 2, \quad \bar{\alpha}_2 = 5, \quad \bar{\beta}_2 = 2$$

also satisfies (8.40) and (8.41) and is thus an M-stationary point. Unfortunately, if we plot the profit functions using this point (see Figure 8.2), we see that this point is merely M-stationary and not a solution. This is due to the fact that M-stationarity conditions, as noted earlier, may accept too many stationary points, some of which are not solutions. Yet the value of the quantifying information obtained from the relations is undeniable and though in the next few examples

Figure 8.2: Profit functions at $(\bar{\alpha}_1, \bar{\beta}_1, \bar{\alpha}_2, \bar{\beta}_2) = (1, 2, 5, 2)$

we will derive sharper conditions via S-stationarity, such an argument is not applicable in many settings.

In order to continue, we need to argue that the loss-less EPEC can be transformed into a similar EPEC that does not take transmission into account near solutions types without congestion.

Example 8.3 (a reduced spot market EPEC). *We will proceed in a generalized setting to that which was considered in the Example 8.1, i.e., we consider the spot market EPEC (3.9) without losses and solutions $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ in which no congestion is allowed ($k = m$) and there may be inactive producers $1 \leq l \leq N$. In this setting, the demand satisfaction relations reduce to*

$$q + By \geq d. \tag{8.43}$$

Since we are disregarding congestion, the demand satisfaction relation may be further reduced to its summarized version $1^T q \geq 1^T d$, where $1^T = (1, \dots, 1)$. In other words, it suffices to require that the total generation of energy meets the total demand. Indeed, the reduced version follows from (8.43) upon left multiplication with 1^T and using statement 1. of Lemma 7.1. Conversely, given the relation $1^T q \geq 1^T d$, there always exists some appropriate flow vector, such that (8.43) holds true. This is a consequence of the well-known Gale-Hoffman inequalities (see, e.g., Rockafellar [1984]) for the special case of unconstrained flows y, which is exactly the case we are considering here due to the absence of congestion around solutions investigated. Note that an analogous statement would be false in the presence of losses. On the other hand, the objective in the ISO problem (3.7) does not depend on the flow y. Therefore, the flow becomes meaningless in the lossless ISO problem and thus for the entire EPEC. Once some $\bar{q}$ is fixed, one may recover a feasible flow $\bar{y}$ afterwards. Summarizing, in the lossless case we are allowed to remove the y-variables from the EPEC (4.37) upon replacing the feasible set G in the ISO problem by a reduced one:

$$\min \left\{ \sum_{i=1}^{N} \alpha_i q_i + \beta_i q_i^2 \,\middle|\, q \in \hat{G} \right\}, \quad \hat{G} := \{ q \in \mathbb{R}^n \,\big|\, q \geq 0,\ 1^T q \geq 1^T d \}. \tag{8.44}$$

Then the lossless EPEC becomes

$$\min_{\substack{(\alpha_i, \beta_i) \in \mathbb{R}^2 \\ q \in \tilde{G}}} \left\{ \hat{f}_i(\alpha, \beta, q) \,\middle|\, 0 \in \hat{F}(\alpha, \beta, q) + N_{\hat{G}}(q) \right\} \quad (i = 1, \dots, N). \tag{8.45}$$

Here $\hat{f}_i(\alpha, \beta, q) := f_i(\alpha, \beta, q, y)$ and $\hat{F}(\alpha, \beta, q) := \alpha + 2[\operatorname{diag} \beta] q$ are the same functions as in the original EPEC (3.9), whereas in the latter, the formal dependence on y is removed.

Now that we have shown that our EPEC can be reduced in the setting considered in Example 8.1, we directly calculate the Fréchet normal cone to the graph of the solution mapping associated with the generalized equation in (8.45). But first, we need need a few structural properties of this reduced EPEC (8.45) in order to return to the example. The following proposition states that LICQ holds at $\bar{q}$ and SSOSC at $(\bar{\alpha}, \bar{\beta}, \bar{q})$. Therefore, we obtain strong regularity of the generalized equation induced by the KKT conditions of (8.44) as well.

As in our earlier considerations, we use a local description of $\tilde{G}$. This allows us to reduce the non-negativity constraints to those components of q belonging to non-active generators. The demand satisfaction inequality has to be included too in this local description, since at a solution it is always satisfied as an equality; a consequence of statement 2. in Lemma 7.2. Therefore, around some $\bar{q}$ as in (7.1), $\tilde{G}$ may be locally described by

$$h(q) \leq 0, \quad h(q) := \left(1^T d - 1^T q, -q^{(2)}\right)^T \tag{8.46}$$

and $q^{(2)} := (q_{l+1}, \ldots, q_N)$.

Proposition 8.3 (structural properties of the reduced EPEC). *Assume $\rho_j = 0$ for $j = 1, \ldots, m$ and let $(\bar{\alpha}, \bar{\beta}, \bar{q})$ be a solution to (8.45) such that $\bar{\alpha}_i, \bar{\beta}_i > 0$ $(i = 1, \ldots, N)$, and $\bar{q}_i > 0$ $(i = 1, \ldots, l)$, $\bar{q}_i = 0$, $(i = l+1, \ldots, N)$, with $1 \leq l \leq N$. Then*

1. *LICQ holds at $\bar{q} \in \hat{G}$*
2. *SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{q})$*

Thus in particular, the generalized equation

$$0 \in \begin{bmatrix} \mathcal{L}(\alpha, \beta, q, \eta, \xi) \\ 1^T q - 1^T d \\ q^{(2)} \end{bmatrix} + N_{\mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^{N-l}}(q, \eta, \xi), \tag{8.47}$$

is strongly regular at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{\eta}, \bar{\xi})$. Here,

$$\mathcal{L}(\alpha, \beta, q, \eta, \xi) := \alpha + 2[\text{diag } \beta]q + \nabla^T h(q)(\eta, \xi)$$

and $(\bar{\eta}, \bar{\xi})$ is the uniquely defined Lagrange multiplier associated with $(\bar{\alpha}, \bar{\beta}, \bar{q})$ in (8.44).

Proof. Following the same argument as in the proof of Proposition 7.2, we end up with

$$\nabla_q \mathcal{L}(\bar{\alpha}, \bar{\beta}, \bar{q}) = 2[\text{diag } \bar{\beta}],$$

which is always positive definite, regardless of $\ker(\nabla h(\bar{q}))$. Whence, statement 1. Furthermore, the Jacobian of the feasible set mapping to (8.44), $\nabla h(\bar{q})$, has the structure:

$$\begin{pmatrix} -1 & \ldots & -1 \\ 0 & | & -I_{N-l} \end{pmatrix}$$

This matrix is always surjective, which incidentally justifies the uniqueness of the multiplier stated in the assertion of this proposition and insures that LICQ holds at $\bar{q} \in \hat{G}$. Therefore, strong regularity follows as in Proposition 7.2. $\square$

Example 8.2 (S-stationarity via direct derivation of $\widehat{N}_{\mathrm{gph}\, S_i}$). Consider a solution $(\bar{\alpha}, \bar{\beta}, \bar{q})$ to the reduced EPEC (8.45) with $N = 2$ satisfying the conditions of Proposition 8.3 under the assumption that $l = 1$. In order to derive strong stationarity conditions for this EPEC, we need to derive two Fréchet normal cones to the graphs of S_i for $i = 1, 2$. However, since the derivations are analogous, we only provide the proof for $i = 1$. Note that $\bar{q}_1 = d$ and $\bar{q}_2 = 0$, where d is the total demand.

The following derivation follows some of the proof of Proposition 4.2 and thus, we will subsequently use part of its notation. Moreover, due to Proposition 8.3, the important arguments from the proof of Proposition 4.2 pertaining to the directional differentiability of the solution mapping still hold. Begin by noting that

$$\nabla_{\alpha_1,\beta_1} \hat{F}(\bar{\alpha}, \bar{\beta}, \bar{q}) = \begin{pmatrix} 1 & 2d \\ 0 & 0 \end{pmatrix}, \quad \nabla_q \hat{F}(\bar{\alpha}, \bar{\beta}, \bar{q}) = \begin{pmatrix} 2\bar{\beta}_1 & 0 \\ 0 & 2\bar{\beta}_2 \end{pmatrix}.$$

Then

$$T_{\mathrm{gph}\, S_1}(\bar{\alpha}_1, \bar{\beta}_1, \bar{q}) = \Phi^{-1}(\Gamma),$$

where

$$\Phi(h, v) = \begin{pmatrix} v_1 \\ v_2 \\ -h_1 - 2dh_2 - 2\bar{\beta}_1 v_1 \\ -2\bar{\beta}_2 v_2 \end{pmatrix}, \quad \Gamma = \mathrm{gph}\, N_K.$$

Here, $K = T_{\hat{G}}(\bar{q}) \cap \{\hat{F}(\bar{\alpha}, \bar{\beta}, \bar{q})\}^{\perp}$. First note that

$$T_{\hat{G}}(\bar{q}) = T_{\hat{G}}(d, 0) = \mathbb{R}_+ \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Using the KKT-conditions from the reduced ISO problem (8.44), we observe

$$\{\hat{F}(\bar{\alpha}, \bar{\beta}, \bar{q})\}^{\perp} = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \supseteq T_{\hat{G}}(\bar{q}).$$

Then it is clear that

$$K = \mathbb{R}_+ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad K^- = \{(w_1^*, w_2^*) \,|\, w_2^* \le w_1^*\}$$

and thus

$$\Gamma = \{(w, w^*) \,|\, w_1 \le 0, w_2 = -w_1, w_2^* \le w_1^*, w_1 w_1^* + w_2 w_2^* = 0\}.$$

After making a case distinction with $w_1 = 0$ and $w_1 < 0$, we can write $\Gamma = \Gamma_1 \cup \Gamma_2$, where

$$\Gamma_1 = \{(w, w^*)\,|w_1 = w_2 = 0, w_2^* \le w_1^*\}$$
$$\Gamma_2 = \{(w, w^*)\,|w_1 < 0, w_2 = -w_1, w_1^* = w_2^*\}$$

Since

$$\bar{\Gamma}_2 = \{(w, w^*)\,|w_1 \le 0, w_2 = -w_1, w_1^* = w_2^*\} \subseteq \Gamma,$$

we may write $\Gamma = \Gamma_1 \cup \bar{\Gamma}_2$, where both sets in the union are then closed. Therefore,

$$T_{\mathrm{gph}\, S_1}(\bar{\alpha}_1, \bar{\beta}_1, \bar{q}) = \Phi^{-1}(\Gamma) = \Lambda_1 \cup \Lambda_2,$$

where

$$\Lambda_1 := \Phi^{-1}(\Gamma_1), \quad \Lambda_2 := \Phi^{-1}(\Lambda_2).$$

Utilizing the definitions of Φ, Γ_1 and Γ_2 results in the following

$$\Lambda_1 = \{(h, v)\,|v_1 = v_2 = 0, h_1 + 2dh_2 \le 0\}$$
$$\Lambda_2 = \left\{(h, v)\,\middle|\,v_1 \le 0, v_2 = -v_1, h_1 + 2dh_2 + 2\bar{\beta}_1 v_1 = 2\bar{\beta}_2 v_2\right\}.$$

Recalling the definition of the Fréchet normal cone, we have

$$\widehat{N}_{\mathrm{gph}\, S_1}(\bar{\alpha}_1, \bar{\beta}_2, \bar{q}) = \left[T_{\mathrm{gph}\, S_1}(\bar{\alpha}_1, \bar{\beta}_1, \bar{q})\right]^- = [\Lambda_1 \cup \Lambda_2]^- = \Lambda_1^- \cap \Lambda_2^-.$$

Furthermore,

$$\Lambda_1^- = \mathbb{R}_+ \begin{pmatrix} 1 \\ 2d \\ 0 \\ 0 \end{pmatrix}, \quad \Lambda_2^- = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mathrm{span}\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2d \\ 2\bar{\beta}_1 \\ -2\bar{\beta}_2 \end{pmatrix}\right\}.$$

Then by substition, we have

$$\widehat{N}_{\mathrm{gph}\, S_1}(\bar{\alpha}_1, \bar{\beta}_2, \bar{q}) = \left\{(x_1^*, x_2^*, y_1^*, y_2^*)\,\middle|\,x_1^* \ge 0, x_2^* = 2dx_1^*. y_1^* \ge y_2^* + 2x_1^*(\bar{\beta}_1 + \bar{\beta}_2)\right\}.$$

Referring back to Definition 4.1, the S-stationarity conditions for the MPEC associated with $i = 1$ read

$$0 \in \nabla_{\alpha_1, \beta_1, q_1, q_2} \hat{f}_1(\bar{\alpha}, \bar{\beta}, \bar{q}) + \widehat{N}_{\mathrm{gph}\, S_1}(\bar{\alpha}_1, \bar{\beta}_2, \bar{q}).$$

Upon substituting the previous relation and proceeding in a manner analogously to that which we used in the earlier examples, we see that the strong stationarity conditions can be reduced to the following, where the reader shall note that (8.49) follows from the derivation of $N_{\mathrm{gph}\, S_2}$:

$$(\bar{\alpha}_1 - \gamma_1) + 2(\bar{\beta}_1 - \bar{\beta}_2 - \delta_1)\bar{q}_1 \ge 0 \tag{8.48}$$
$$\bar{\alpha}_2 \le \gamma_2 \tag{8.49}$$
$$\bar{\alpha}_1 + 2\bar{\beta}_1 \bar{q}_1 = \bar{\alpha}_2 \tag{8.50}$$

Notice that of the two different solution and parameter constellations considered in Example 8.1 only the first combination fulfills (8.49). Therefore, had we known this extra information, then we would have been immediately able to say that the combination considered for Figure 8.2 cannot be a solution, since in that setting $\bar{\alpha}_2 > \gamma_2$. Moreover, we can use this new information to derive the relation

$$\gamma_2 > \gamma_1 - 2\delta_1 d$$

This last inequality states that no EPEC solution of the type considered is possible in the event $\gamma_2 \leq \gamma_1 - 2\delta_1 d$, thus providing us with a glimpse into how could use stationarity conditions for determining the sensitivity of certain EPECs on their parameters.

Though we observed in Examples 8.2 and 8.2 that S-stationarity conditions may be more selective than M-stationarity conditions, the direct derivation of the Fréchet normal cones in Example 8.2 is in general not applicable, even for a three player setting. This is rather unfortunate, however, the next Example shows (via Theorem 4.3) that we can still find cases in which M-stationarity conditions can be used to obtain S-stationarity conditions, without the need to reduce the EPEC.

Example 8.3 (Obtaining S-stationarity via M-stationarity). Let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution to (3.9) satisfying (7.1) such that

$$N = l = 2, \quad m - k = 1, \quad \rho = 0, \quad B = (-1, 1)^T,$$

that is, $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ is a solution such that both producers are active, the line is congested, losses are ignored, and power is flowing from node 1 to node 2. In such a case, we can refer to Lemma 7.2 statement 1(b), which states that LICQ holds at $(\bar{q}, \bar{y}) \in G$ and Proposition 7.2, which indicates that SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$. Thus we can apply Theorem 8.1. In the current setting (8.3)-(8.8) reduce to

$$(v^i)_i = \bar{q}_i > 0, \ (i = 1, 2) \tag{8.51}$$
$$(\bar{\alpha}_i - \gamma_i) + 2(\bar{\beta}_i - \delta_i)\bar{q}_i = -(w^i)_i, \ (i = 1, 2) \tag{8.52}$$
$$2\bar{\beta}_j (v^i)_j = (w^i)_j, \ (j \neq i, j = 1, 2) \tag{8.53}$$
$$(w^i)_1 = (w^i)_2, \ (i = 1, 2) \tag{8.54}$$
$$(v^i)_1 = -(v^i)_2 = (v^i)_3, \ (i = 1, 2) \tag{8.55}$$
$$\bar{\alpha}_1 + 2\bar{\beta}_1 \bar{q}_1 = \bar{\alpha}_2 + 2\bar{\beta}_2 \bar{q}_2 \tag{8.56}$$

Note that (8.55) follows from (8.8), where we as in the previous examples, we can demonstrate that $I_+ = \{1, 2\}$ and $I_0 = \{3\}$.

In addition, we have for all $i = 1, 2$ that $\nabla H_{I_0}(\bar{q}, \bar{y}) v^i = \nabla H_3(\bar{q}, \bar{y}) v^i > 0$. That is, there exists an M-stationarity multiplier such that assumption 4. from Theorem 4.3 is fulfilled, in which case we see that conditions (8.51)-(8.56) are not just M, but S-stationarity conditions as well[2]. As before, we may reduce relations (8.51)-(8.56), for which we obtain

$$(\bar{\alpha}_1 - \gamma_1) + 2(\bar{\beta}_1 - \bar{\beta}_2 - \delta_1)\bar{q}_1 = 0 \tag{8.57}$$

$$(\bar{\alpha}_2 - \gamma_2) + 2(\bar{\beta}_2 - \bar{\beta}_1 - \delta_2)\bar{q}_2 = 0. \tag{8.58}$$

Given $H(\bar{q}, \bar{y}) = 0$ (by Lemma 7.2 statement 2), $\bar{q}_2 = d - \bar{q}_1$, where $d = d_1 + d_2$ is the total demand in the network. Thus, we can use (8.56) to solve for the production levels:

$$\bar{q}_1 = \frac{\bar{\alpha}_2 - \bar{\alpha}_1 + 2\bar{\beta}_2 d}{2(\bar{\beta}_1 + \bar{\beta}_2)}, \quad \bar{q}_2 = d - \frac{\bar{\alpha}_2 - \bar{\alpha}_1 + 2\bar{\beta}_2 d}{2(\bar{\beta}_1 + \bar{\beta}_2)}. \tag{8.59}$$

Moreover, if we let $\gamma_1 = 4$, $\gamma_2 = 1$, $\delta_1 = 1$, $\delta_2 = 2$, and $d = 1$; then we see that $\bar{\alpha}_1 = 5$, $\bar{\beta}_1 = 1$, $\bar{\alpha}_2 = 5/2$, $\bar{\beta}_2 = 2$, fulfills (8.57) and (8.58). Plotting the profit functions, we see that this point is not only an S-stationary point, but also a solution; refer to Figure 8.3.

Figure 8.3: Profit functions at $(\bar{\alpha}_1, \bar{\beta}_1, \bar{\alpha}_2, \bar{\beta}_2) = (5, 1, 5/2, 2)$

Again, generator 1's profit function is on the left and generator 2' profit function is on the right. Therefore, we have found an example in which strict complementarity fails and the fourth condition of Theorem 4.3 holds. Thus demonstrating the usefulness of the new condition

8.3 Explicit M-stationarity Conditions for the Spot Market SEPEC

In this last section, we consider the stochastic spot market EPEC. In the following, we will not be considering cases where the transmission lines can become congested, therefore, the reader should not confuse the index k with the previous usage of k for the number of uncongested tranmission lines. The main purpose of this section is to demonstrate that explicit M-stationarity conditions can also be derived in a stochastic setting. Assuming a discrete distribution of random demands, i.e., d is a random vector with realizations $\{d^k\}_{k=1}^K$ with scenario-probabilities p_k, $k = 1, \ldots, K$, then the (loss-less) SEPEC becomes

$$\min_{\substack{(\alpha_i, \beta_i) \in \mathbb{R}^2 \\ \mathbf{q}, \mathbf{y}}} \Big\{ -\tilde{f}_i(\alpha_i, \beta_i, \mathbf{q}, \mathbf{y}) \;\Big|\; 0 \in F^k(\alpha, \beta, q^k, y^k) + N_{G^k}(q^k, y^k),\ (k = 1, \ldots, K) \Big\} \quad (i = 1, \ldots, N). \tag{8.60}$$

[2]That is, when considering the M-stationarity conditions derived for each of the MPECs composing the EPEC, there exist the needed multipliers for $i = 1, 2$.

Here, $\mathbf{q} = (q^1, \dots, q^K)$ and $\mathbf{y} = (y^1, \dots, y^K)$ and

$$\begin{aligned}
\tilde{f}_i(\alpha_i, \beta_i, \mathbf{q}, \mathbf{y}) &:= \sum_{k=1}^{K} p_k[(\alpha_i - \gamma_i)q_i^k + (2\beta_i - \delta_i)(q_i^k)^2] \\
F^k(\alpha, \beta, q^k, y^k) &:= \begin{pmatrix} \alpha + 2[\operatorname{diag}\beta]q^k \\ 0 \end{pmatrix} \\
G^k &:= \left\{ (q, y) \in \mathbb{R}^{N+m} \;\middle|\; \begin{array}{c} q^k + By^k \geq d^k \\ 0 \leq q^k \leq \hat{q} \\ -\hat{y} \leq y^k \leq \hat{y} \end{array} \right\}.
\end{aligned}$$

By defining

$$\widetilde{F}(\alpha, \beta, \mathbf{q}, \mathbf{y}) := \begin{pmatrix} F^1(\alpha, \beta, q^1, y^1) \\ \vdots \\ F^K(\alpha, \beta, q^K, y^K) \end{pmatrix}, \;\; \mathbf{G} := G^1 \times \cdots \times G^k \times \cdots \times G^K,$$

using elementary properties of the normal cone to closed convex sets allows us to rewrite (3.15) as a standard EPEC

$$\min_{\substack{(\alpha_i, \beta_i) \in \mathbb{R}^2 \\ \mathbf{q}, \mathbf{y}}} \left\{ -\tilde{f}_i(\alpha_i, \beta_i, \mathbf{q}, \mathbf{y}) \,\middle|\, 0 \in \widetilde{F}(\alpha, \beta, \mathbf{q}, \mathbf{y}) + N_{\mathbf{G}}(\mathbf{q}, \mathbf{y}), \right\} \quad (i = 1, \dots, N). \tag{8.61}$$

Nevertheless, we still need to show that some constraint qualifications hold and that the perturbation mappings

$$\Psi_i(\mathbf{u}) := \left\{ (\alpha_i, \beta_i, \mathbf{q}, \mathbf{y}) \in \mathbb{R}^{2N} \times \mathbb{R}^{(N+m)k} \,\middle|\, \mathbf{u} \in \widetilde{F}(\bar{\alpha}_{-i}, \alpha_i, \bar{\beta}_{-i}, \beta_i, \mathbf{q}, \mathbf{y}) + N_{\mathbf{G}}(\mathbf{q}, \mathbf{y}) \right\}$$

are calm at $(0, \bar{\alpha}_i, \bar{\beta}_i, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ for all $i = 1, \dots, N$, where $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ is a solution to (8.61). We start with the following lemma.

Lemma 8.1 (**Structural Properties of the Stochastic ISO Problem**). *Suppose $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ is a solution to (8.61) such that*

1. $\bar{\alpha}_i, \bar{\beta}_i > 0 \quad (i = 1, \dots, N)$
2. $\rho_j = 0 \quad j = 1, \dots, m$ *(no losses)*
3. $-\hat{y}_j < \bar{y}_j^k < \hat{y}_j \quad (j = 1, \dots, m;\ k = 1, \dots, K)$ *(no congestion almost surely)*
4. *B generates an acyclic network*

Then

1. *LICQ holds at $(\bar{\mathbf{q}}, \bar{\mathbf{y}}) \in \mathbf{G}$*
2. *Strict complementarity holds for each k-th set of demand satisfaction constraints $(k = 1, \dots, K)$*

3. $\widetilde{H}(\mathbf{q}, \mathbf{y}) = 0$*.*

4. SSOSC holds at $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$

We let $\mathbf{l} := \sum_{k=1}^{K} |l_k|$*, where* l_k *is the number of active producers at a solution in scenario* k*,* $k = 1, \ldots, K$*.*

Proof. Let $\widetilde{H} \in \mathbb{R}^{(N+m)k} \to \mathbb{R}^{2NK-\mathbf{l}}$ be the constraint mapping locally describing $\mathbf{G}$, i.e,

$$\widetilde{H}(\mathbf{q}, \mathbf{y}) = \begin{pmatrix} H(q^1, y^1) \\ \vdots \\ H(q^K, y^K) \end{pmatrix} = \begin{pmatrix} d^1 - q^1 - By^1 \\ -q^{1,(2)} \\ \vdots \\ d^K - q^K - By^K \\ -q^{K,(2)} \end{pmatrix}$$

Then

$$\nabla_{\mathbf{q},\mathbf{y}} \widetilde{H}(\bar{\mathbf{q}}, \bar{\mathbf{y}}) = \begin{pmatrix} \nabla_{q^1,y^1} H(\bar{q}^1, \bar{y}^1) & & 0 \\ & \ddots & \\ 0 & & \nabla_{q^K,y^K} H(\bar{q}^K, \bar{y}^K) \end{pmatrix},$$

where for any $k \in \{1, \ldots, K\}$

$$\nabla_{q^1,y^1} H(\bar{q}^1, \bar{y}^1) = \begin{pmatrix} -I_1^k & 0 & -B^{k,1} \\ 0 & -I_2^k & -B^{k,2} \\ 0 & -I_2^k & 0 \end{pmatrix}.$$

Here, the superscript k is to remind the reader that it is possible in each of the considered scenarios for the amount of active and inactive producers to be different. Nevertheless, using the same argument as in Lemma 7.2, we can show that each of these submatrices is surjective for all $k = 1, \ldots, K$ and therefore, $\nabla_{\mathbf{q},\mathbf{y}} \widetilde{H}(\bar{\mathbf{q}}, \bar{\mathbf{y}})$ as well. Hence, LICQ holds at $(\bar{\mathbf{q}}, \bar{\mathbf{y}}) \in \mathbf{G}$.

Following similarly to the previous argument, we can again use the proof of Lemma 7.2 to show that statement 2 holds. Indeed, since $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ is a solution to (8.61), it is also a solution to the generalized equation in (8.61). This, along with the surjectivity of $\nabla_{\mathbf{q},\mathbf{y}} \widetilde{H}(\bar{\mathbf{q}}, \bar{\mathbf{y}})$, allows us to write for each $k = 1, \ldots, K$, $\exists \bar{\lambda}^k \in \mathbb{R}_+^{2N-l_k}$ such that

$$0 = \begin{pmatrix} \alpha + 2[\operatorname{diag} \beta] q^k \\ 0 \end{pmatrix} + \nabla_{q^k,y^k}^T H(\bar{q}^k, \bar{y}^k) \bar{\lambda}^k,$$

The rest of the argument follows exactly that which was used in the proof of Lemma 7.2.

Using statement 2., we can follow as in the proof of Lemma 7.2 to show that each component of $\widetilde{H}$, i.e., $H(q^k, y^k)$ equals zero are a solution. Whence, statement 3.

Finally, we observe that in the current setting, the Jacobian of the vector

Lagrangian with respect to $(\mathbf{q}, \mathbf{y})$ becomes

$$\nabla_{\mathbf{q},\mathbf{y}}\mathcal{L}(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}}, \bar{\lambda}) = \begin{pmatrix} \nabla_{q^1,y^1}\mathcal{L}(\bar{\alpha}, \bar{\beta}, \bar{q}^1, \bar{y}^1, \bar{\lambda}^1) & & 0 \\ & \ddots & \\ 0 & & \nabla_{q^K,y^K}\mathcal{L}(\bar{\alpha}, \bar{\beta}, \bar{q}^K, \bar{y}^K, \bar{\lambda}^K) \end{pmatrix},$$

Here, we use $\bar{\lambda} = (\bar{\lambda}^1, \ldots, \bar{\lambda}^K)$ to denote the associated Lagrange multiplier. The argument demonstrating the positive-definiteness on $\ker\{\nabla_{\mathbf{q},\mathbf{y}}\widetilde{H}(\bar{\mathbf{q}}, \bar{\mathbf{y}})$ analogously follows the proof used in Proposition 7.2. Hence, statement 4. holds. □

Now that we have the necessary constraint qualifications, we can provide the next result. Notice how it would have presented us with relatively no new difficulties, if we had chosen to consider a SEPEC in which transmission losses were included. Nevertheless, we have chosen the lossless case as it is better for illustration.

Proposition 8.4 (strong regularity/Aubin property). *Let $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ be a solution to (8.61) such that*

1. $\bar{\alpha}_i, \bar{\beta}_i > 0 \quad (i = 1, \ldots, N)$
2. $\rho_j = 0 \quad j = 1, \ldots, m$ *(no losses)*
3. $-\hat{y}_j < \bar{y}_j^k < \hat{y}_j$ $(j = 1, \ldots, m)$, $(k = 1, \ldots, K)$ *(no congestion almost surely)*
4. *B generates an acyclic network*

Then

1. *The generalized equation (8.61) is strongly regular at $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$.*
2. *The perturbation mappings Ψ_i associated with the generalized equations in (8.61) have the Aubin property at $(0, \bar{\alpha}_i, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ for all $i = 1, \ldots, N$.*

Proof. By Lemma 8.1, LICQ and SSOSC hold. Then assertions 1. and 2. follow from Proposition 5.1 and Corollary 5.1, respectively. □

Given Lemma 8.1 and Proposition 8.4, we now have the following result providing explicit M-stationarity conditions for the SEPEC (8.61).

Proposition 8.5. *Let $(\bar{\alpha}, \bar{\beta}, \bar{q}, \bar{y})$ be a solution to the SEPEC (8.61) satisfying*

1. $\bar{\alpha}_i, \bar{\beta}_i > 0 \quad (i = 1, \ldots, N)$
2. $\rho_j = 0 \quad j = 1, \ldots, m$ *(no losses)*
3. $-\hat{y}_j < \bar{y}_j^k < \hat{y}_j \quad (j = 1, \ldots, m;\ k = 1, \ldots, K)$ *(no congestion almost surely)*

Then, for $k = 1, \dots, K$ *and* $i = 1, \dots, N$ *there exist* $\bar{\lambda}^k \in \mathbb{R}_+^{2N-l_k}$ *and* $\left(v^{i,k}, w^{i,k}\right) \in \mathbb{R}^{N+m} \times \mathbb{R}^{2N-l_k}$ *such that*

$$\sum_{k=1}^{K} (v^{i,k})_i = \mathbb{E}\bar{q}_i \tag{8.62}$$

$$\sum_{k=1}^{K} (v^{i,k})_i \bar{q}_i^k = \mathbb{E}(\bar{q}_i)^2 \tag{8.63}$$

For $v_a^{i,k} := \left(v_1^{i,k}, \dots, v_N^{i,k}\right)$ *and* $v_b^{i,k} := \left(v_{N+1}^{i,k}, \dots, v_{N+m}^{i,k}\right)$

$$v_a^{i,k} + Bv_b^{i,k} = 0 \tag{8.64}$$

$$(v^{i,k})_{i_j} = 0 \quad \forall j \in \{l_k + 1, \dots, N\} : \bar{\lambda}^k_{N-l_k+j} > 0. \tag{8.65}$$

For $j \in \{1, \dots, l_k\}$

$$(w^{i,k})_{i_j} = \begin{cases} p_k[\gamma_{i_j} - \bar{\alpha}_{i_j} + 2(\delta_{i_j} - 2\bar{\beta}_{i_j})\bar{q}^k_{i_j}] + 2\bar{\beta}_{i_j}(v^{i,k})_{i_j} & \text{if } i_j = i \\ 2\bar{\beta}_{i_j}(v^{i,k})_{i_j} & \text{if } i_j \neq i \end{cases} \tag{8.66}$$

and for $j \in \{l_k + 1, \dots, N\}$

$$(w^{i,k})_{i_j} + (w^{i,k})_{N-l_k+j} = \begin{cases} p_k[\gamma_{i_j} - \bar{\alpha}_{i_j}] + 2\bar{\beta}_{i_j}(v^{i,k})_{i_j} & \text{if } i_j = i \\ 2\bar{\beta}_{i_j}(v^{i,k})_{i_j} & \text{if } i_j \neq i \end{cases} \tag{8.67}$$

If $j \in \{l_k + 1, \dots, N\}$*, then*

$$(w^{i,k})_{N+j-l_k} = 0 \quad \text{if } \bar{\lambda}^k_{N+j-l_k} = 0,\ (v^{i,k})_{i_j} > 0 \tag{8.68}$$

$$(w^{i,k})_{N+j-l_k} \geq 0 \quad \text{if } \bar{\lambda}^k_{N+j-l_k} = 0,\ (v^{i,k})_{i_j} < 0, \tag{8.69}$$

and the KKT-conditions reduce to

$$B^T w^{i,k} = 0 \tag{8.70}$$

$$\bar{\alpha}_{i_1} + 2\bar{\beta}_{i_1}\bar{q}^k_{i_1} = \bar{\alpha}_{i_j} + 2\bar{\beta}_{i_j}\bar{q}^k_{i_j} \quad (j = 1, \dots, l_k) \tag{8.71}$$

$$= \bar{\alpha}_{i_{j'}} - \bar{\lambda}^k_{N+j'-l_k} \quad (j' = l_k + 1, \dots, N) \tag{8.72}$$

Proof. These relations are the direct result obtained by applying Proposition 8.2 for each scenario; (8.62) and (8.64) result in summing over the relations $(v^{i,k})_i = p_k\bar{q}_i^k$ and $(v^{i,k})_i\bar{q}_i^k = p_k(\bar{q}_i^k)^2$ with respect to k. □

We end this chapter with an interesting result demonstrating that for a two player setting at an equilibrium, both generators are participating with positive probability.

Proposition 8.6. *Let* $N = 2, m = k = 1$ *and consider a solution* $(\bar{\alpha}, \bar{\beta}, \bar{\mathbf{q}}, \bar{\mathbf{y}})$ *of the SEPEC (8.61) satisfying the assumptions of Theorem 8.5. Then, both producers are active with positive probability.*

Proof. Assume the statement of the theorem is false. Then, without loss of generality, the second producer becomes inactive almost surely. In other words,

$$\bar{q}_2^k = 0 \quad \forall k \in \{1, \ldots, K\} \tag{8.73}$$

Then $l_k = 1$ for all k and $i_1 = 1, i_2 = 2$ in the stationarity conditions. Moreover, $\bar{q}_1^k = d^k$ for all k. From (8.64), we infer that

$$(v^{1,k})_1 = -(v^{1,k})_2 = -(v^{1,k})_3.$$

Moreover, (8.65) yields that $\bar{\lambda}_3^k = 0$, whenever $(v^{1,k})_2 \neq 0$, hence

$$\bar{\lambda}_3^k = 0 \quad \forall k \in \{1, \ldots, K\} : (v^{1,k})_1 \neq 0 \tag{8.74}$$

From (8.72) and $\bar{q}_1^k = d^k$ it follows that

$$\bar{\alpha}_1 + 2\bar{\beta}_1 d^k = \bar{\alpha}_2 - \bar{\lambda}_3^k \quad \forall k \in \{1, \ldots, K\}.$$

Consequently, since $\bar{\lambda}_3^k \geq 0$ for all k, one has that

$$d^k \leq \frac{\bar{\alpha}_2 - \bar{\alpha}_1}{2\bar{\beta}_1} \quad \forall k \in \{1, \ldots, K\}.$$

Also, (8.74) implies

$$d^k = \frac{\bar{\alpha}_2 - \bar{\alpha}_1}{2\bar{\beta}_1} \quad \forall k \in \{1, \ldots, K\} : (v^{1,k})_1 \neq 0.$$

Let $k^* \in \{1, \ldots, K\}$ be the unique[3] scenario such that $d^k \leq d^{k^*}$ for all k. Then, $d^k < d^{k^*}$ for all $k \neq k^*$ and, consequently,

$$(v^{1,k^*})_1 \neq 0 \quad \text{and} \quad (v^{1,k})_1 = 0 \quad \forall k \in \{1, \ldots, K\} : k \neq k^*.$$

Then, by (8.62) and $\bar{q}_1^k = d^k$, we arrive at $(v^{1,k^*})_1 = \mathbb{E}d$. Now, with (8.63) with establish the contradiction

$$\begin{aligned} d^{k^*}\mathbb{E}d = d^{k^*}(v^{1,k^*})_1 &= \sum_{k=1}^{K} (v^{1,k})_1 d^k = \mathbb{E}(\bar{q}_1)^2 = \sum_{k=1}^{K} p_k (\bar{q}_1^k)^2 = \sum_{k=1}^{K} p_k (d^k)^2 \\ &< \sum_{k=1}^{K} p_k d^k d^{k^*} = d^{k^*} \sum_{k=1}^{K} p_k d^k = d^{k^*}\mathbb{E}d, \end{aligned}$$

which holds true, whenever $K \geq 2$. Consequently, our assumption (8.73) was wrong and the Theorem is proved. □

Note how we used the stationarity conditions in the *reverse* direction to prove

[3] The holds without loss of generality. Indeed, in the event there exist more than one scenario in which the total demand realizations are equivalent, we can combine these scenarios into a single scenario k^*, where probability of this scenario p_{k^*} is obtained by summing over the respective scenarios involved.

the result of Proposition 8.6, thereby further demonstrating the usefulness of explicit necessary stationarity conditions.

Abbreviations

LICQ	Linear Independence Constraint Qualification
MFCQ	Mangasarian-Fromowitz Constraint Qualification
CRCQ	Constant Rank Constraint Qualification
SSOSC	Strong Second-Order Sufficient Condition
KKT	Karush-Kuhn Tucker
MPEC	Mathematical Program with Equilibrium Constraints
EPEC	Equilibrium Problem with Equilibrium Constraints
ISO	Independent System Operator
SMPEC	Stochastic Mathematical Program with Equilibrium Constraints
SEPEC	Stochastic Equilibrium Problem with Equilibrium Constraints

Bibliography

J.-P. Aubin. Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions. In L. Nachbin, editor, *Mathematical Analysis and Applications*, pages 159–229. Academic Press, New York, 1981.

J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.

B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-Linear Parametric Optimization*. Akademie-Verlag, Berlin, 1982.

M. R. Baye, G. Tian, and J. Zhou. Characterizations of the existence of equilibria in games with discontinuous and non-quasiconcave payoffs. *The Review of Economic Studies*, 60(4):935–948, 1993.

P. Beremlijski, J. Haslinger, M. Kočvara, R. Kučera, and J. V. Outrata. Shape optimization in three-dimensional contact problems with coulomb friction. *SIAM Journal on Optimization*, 20:416–444, 2009.

P. Beremlijski, J. Haslinger, M. Kočvara, and J. V. Outrata. Shape optimization in contact problems with coulomb friction. *SIAM Journal on Optimization*, 13: 561–587, 2002.

C. A. Berry, B. F. Hobbs, W. A. Meroney, R. P. OŃeill, and W. R. Stewart. Understanding how market power can arise in network competition: a game theoretic approach. *Utilities Policy*, 8:139–158, 1999.

N. Biggs. *Algebraic Graph Theory*. Cambridge Univeristy Press, Cambridge, 2 edition, 1994.

J. R. Birge and F. V. Louveaux. *Introduction to Stochastic Programming*. Springer-Verlag, Berlin, 1997.

J.-F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, Berlin, 2000.

S. Borenstein, J. Bushnell, and S. Stoft. The competitive effects of transmission capacity in a deregulated electricity industry. *RAND Journal of Economics*, 31 (2):294–325, 2000.

C. Castaing and M. Valadier. *Convex Analysis and Measurable Multifunctions*. Springer-Verlag, Berlin, 1977.

A. A. Cournot. Recherches sur les principes mathematiques de la theorie des richesses. In *Libraire de Sciences Politiques et Sociale*. M. Riviere & Cie, Paris, 1838.

G. Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences of the U.S.A.*, 38:386–393, 1952.

V. DeMiguel and H. Xu. A stochastic multiple-leader stackelberg model: Analysis, computation, and application. *Operations Research*, 2009. Published online in *Articles in Advance* DOI: 10.1287/opre.1080.0686.

S. Dempe. *Foundations Of Bilevel Programming*. Kluwer Academic Publishers, Dodrecht, 2002.

S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52:333–359, 2003.

A. L. Dontchev and R. T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM Journal on Optimization*, 7:1087–1105, 1996.

A. Ehrenmann. *Equilibrium problems with equilibrium constraints and their application to electricity markets*. PhD thesis, Fitzwilliam College, Cambridge University, 2004a.

A. Ehrenmann. Manifolds of multi-leader cournot equilibria. *Operations Research Letters*, 32:121–125, 2004b.

J. F. Escobar and A. Jofre. Oligopolistic competition in electricity spot markets. Available at http://ssrn.com/abstract=878762, December 2005.

A. Evgrafov and M. Patriksson. On the existence of solutions to stochastic mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 121(1):65–76, 2004.

F. Facchinei, H. Jiang, and L. Qi. A smoothing method for mathematical programs with equilibrium constraints, 1996.

F. Facchinei and J.-S. Pang. *Finite-dimensional Variational Inequalities and Complementarity Problems Vol. I*. Springer-Verlag, New York, 2003.

K. Fan. Some properties of convex sets related to fixed point theorems. *Mathematical Annals*, 266:519–537, 1984.

A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley-Interscience, New York, 1968.

M. L. Flegel, C. Kanzow, and J. V. Outrata. Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints. *Set-Valued Analysis*, 15(2):139–162, 2007.

J. W. Friedman. *Oligopoly and the Theory of Games*. North-Holland, Amsterdam, 1977.

L. Hakonsen. *Essays on Taxation, Efficiency and the Environment*. PhD thesis, Norwegian School of Economics & Business Administration, April 1998.

P. T. Harker. A variational inequality approach for the determination of oligopolistic market equilibrium. *Mathematical Programming*, 30:105–111, 1984.

P.T. Harker and J.-S. Pang. Finite-dimensional variational inequalities and complementarity problems: a survey or theory, algorithms, and applications. *Mathematical Programming*, 60:161–220, 1990.

J. Heerda and B. Kummer. Characterization of calmness for banach space mappings. Technical report, Humbodlt-Universitaet-zu-Berlin, November 2006. HUB Preprint Reihe (Mathematik).

R. Henrion. On constraint qualifications. *Journal of Optimization Theory and Applications*, 72:187–196, 1992.

R. Henrion, A. Jourani, and J. V. Outrata. On the calmness of a class of multifunctions. *SIAM Journal on Optimization*, 13:603–618, 2002.

R. Henrion, B.S. Mordukhovich, and N.M. Nam. Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities. Technical report, Weierstrass Institute Berlin, 2009a. Preprint No. 1399.

R. Henrion and J. V. Outrata. Calmness of constraint systems with applications. *Mathematical Programming*, 104:437–464, 2005.

R. Henrion, J.V. Outrata, and T. Surowiec. M. DFG Research Center MATHEON Preprint: Available at, July 2009b.

R. Henrion, J.V. Outrata, and T. Surowiec. On the co-derivative of normal cone mappings to inequality systems. *Nonlinear Analysis*, 71:1213–1226, 2009c.

R. Henrion, J.V. Outrata, and T. Surowiec. Strong stationary solutions to equilibrium to equilibrium problems with equilibrium constraints with applications to an electricity spot market model. Weierstraß-Institute of Applied Analysis and Stochastics Preprint No. 1396. Available at http://www.wias-berlin.de/publications/preprints/, January 2009d.

R. Henrion and W. Römisch. On m-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. *Applications of Mathematics*, 52:473–494, 2007.

X. Hu. *Mathematical Programs with Complementarity Constraints and Game Theory Modles in Electricity Markets*. PhD thesis, Department of Mathematics and Statistics, University of Melbourne, 2002.

X. Hu and D. Ralph. Using epecs to model bilevel games in restructured electricity markets with locational prices. Technical report, Judge Business School, December 2005. Available at http://www.optimization-online.org/.

X. Hu, D. Ralph, E. K. Ralph, P. Bardsley, and M. C. Ferris. Electricity generation with looped transmission networks: bidding to an ISO. Technical report, Judge Business School, February 2007. Available at http://papers.ssrn.com/sol3/DisplayAbstractSearch.cfm.

A.D. Ioffe and J.V. Outrata. On metric and calmness qualification conditions in subdifferential calculus. *Set-Valued Analysis*, 16:199–227, 2008.

R. Janin. Directional derivative of the marginal function in nonlinear programming. *Mathematical Programming Studies*, 21:110–126, 1984.

D. Klatte and B. Kummer. Generalized kojima-functions and lipschitz stability of critical points. *Computational Optimization and Applications*, 13:61–85, 1999.

D. Klatte and B. Kummer. Constrained minima and lipschitzian penalties in metric spaces. *SIAM Journal on Optimization*, 13(2):619–633, 2002a.

D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization.* Kluwer Academic Publishers, Dodrecht, 2002b.

D. Klatte and B. Kummer. Second-order characterizations of lipschitz stability in nonlinear porgramming. *Journal of Mathematical Sciences*, 116(3):3231–3252, 2003.

D. Klatte and B. Kummer. Optimization methods and stability of inclusions in banach spaces. *Mathematical Programming*, 117:305–330, 2009.

S. Leyffer and T. Munson. Solving multi-leader-common-follower games. Preprint ANL/MCS-P1243-0405, Argonne National Laboratory, Mathematics and Computer Science Division, April 2005.

P. Loridan and J. Morgan. A theoretical approximation scheme for stackelberg problems. *Journal of Optimization Theory and Applications*, 61:95–110, 1989.

S. Lu. Implications of the constant rank constraint qualification. To appear in *Mathematical Programming Series A*, 2009.

Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, Cambridge, 1997.

Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.-Q. Wu. Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints. *Mathematical Programming*, pages 19–76, 1996.

L. Minchenko and S. Stakhovski. On generalized constant rank regularity condition in mathematical programming. To appear in Optimization, 2009.

B. S. Mordukhovich. Maximum principle in problems of time optimal control with nonsmooth constraints. *Journal of Applied Mathematics and Mechanics*, 40:960–969, 1976.

B. S. Mordukhovich. Metric approximations and necessary optimality conditions for general classes of extremal problems. *Soviet Mathematics Doklady*, 22:526–530, 1980.

B. S. Mordukhovich. Complete characterizations of openess, metric regularity, and lipschitzian properties of multifunctions. *Transactions of the Americaion Mathematical Society*, 340:1–36, 1993.

B. S. Mordukhovich. Equilibrium problems with equilibrium constraints via multiobjective optimization. *Optimization Methods and Software*, 19:479–492, 2004.

B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory*. Springer-Verlag, Berlin, 2006a.

B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation. Vol. 2: Applications*. Springer-Verlag, Berlin, 2006b.

B. S. Mordukhovich and J. V. Outrata. Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. *SIAM J. on Optimization*, 18(2), 2007.

B. S. Mordukhovich, J. V. Outrata, and M. Cervinka. Equilibrium problems with complementarity constraints: Case study with applications to oligopolistic markets. *Optimization*, 56:479–494, 2007.

B. S. Mordukhovich and J.V. Outrata. On second-order subdifferentials and their applications. *SIAM Journal on Optimization*, 12(1):139–169, 2001.

O. Morgenstern and J. von Neumann. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

J. F. Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951.

K. Nishimura and J. Friedman. Existence of nash equilibrium in n person games without quasi-concavity. *International Economic Review*, 22:637–648, 1981.

M. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press, Cambridge, Massachusetts, London, England, 1994.

J. V. Outrata. Optimality conditions for a class of mathematical programs with equilibrium constraints. *Mathematics of Operations Research*, 25(3):627–644, 1999.

J. V. Outrata. A generalized mathematical program with equilibrium constraints. *SIAM Journal of Control and Optimization*, 38(5):1623–1638, 2000.

J. V. Outrata. On constraint qualifications for mathematical programs with mixed complementarity constraints. In M. C. Ferris, O. L. Mangasarian, and J.-S. Pang, editors, *Complementarity: Applications, Algorithms and Extensions (Applied Optimization)*, pages 253–272. Springer-Verlag, Berlin, 2001.

J. V. Outrata. A note on a class of equilibrium problems with equilibrium constraints. *Kybernetika*, 40:585–594, 2004.

J. V. Outrata. (personal communication), 2008.

J. V. Outrata, M. Kočvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publishers, Dodrecht, 1998.

J. V. Outrata and M. Červinka. On the implicit programming approach in a class of mathematical programs with equilibrium constraints. Submitted, 2009.

J.-S. Pang and M. Fukushima. Complementarity constraint qualifications and cimplified b-stationarity conditions for mathematical programs with equilibrium constraints. *Computational Optimization and Applications*, 13:111–136, 1999.

J.-S. Pang and M. Fukushima. Quasi-variational inequalities, generalized nash equilibria, and multi-leader-follower games. *Computational Management Science*, 2:21–56, 2005.

M. Patriksson and L. Wynter. Stochastic mathematical programs with equilibrium constraints. *Operations Research Letters*, 25(4):159–167, 1999.

D. Ralph and S. Dempe. Directional derivatives of the solution of a parametric nonlinear program. *Mathematical Programming*, 70:159–172, 1995.

D. Ralph and Y. Smeers. Epecs as models for electricity markets. Invited Paper, 2006 Power Systems Conference and Exposition (PSCE), 2006.

S. M. Robinson. Generalized equations and their solutions, part I: basic theory. *Mathematical Programming Study*, 10:128–141, 1979.

S. M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5: 43–62, 1980.

S. M. Robinson. Some continuity properties of polyhedral multifunctions. *Mathematical Programming Study*, 14:206–214, 1981.

S. M. Robinson. Implicit b-differentiability in generalized equations. Technical report, Mathematics Research Center, University of Wisconsin, 1985. Technical Summary Report No. 2854.

R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer-Verlag, Berlin, 1998.

R.T. Rockafellar. *Network Flows and Monotropic Optimization.* Wiley-Interscience, New York, 1984.

A. Ruszczyński and A. Shapiro, editors. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science.* Elsevier, 2003.

H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.

A. Shapiro. Stochastic programming with equilibrium constraints. *Journal of Optimization Theory and Applications*, 128(1):223–243, 2006.

A. Shapiro and H. Xu. Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. *Optimization*, 57(3): 395–418, 2008.

H. D. Sherali. A multiple leader stackelberg model and analysis. *Operations Research*, 32:309–404, 1984.

H. D. Sherali, A. L. Soyster, and F. H. Murphy. A mathematical programming approach for determining oligopolistic market equilibrium. *Mathematical Programming*, 24:92–106, 1982.

H. D. Sherali, A. L. Soyster, and F. H. Murphy. Stackelberg-nash-cournot equilibria: characterization and computations. *Operations Research*, 31:253–276, 1983.

W. Song. Calmness and error bounds for convex constraint systems. *SIAM Journal on Optimization*, 13:353–371, 2006.

C.-L. Su. *Equilibrium problems with equilibrium constraints: stationarities, algorithms, and applications.* PhD thesis, Stanford University, 2005.

M. Červinka. *Hierarchal Structures in Equilibrium Problems.* PhD thesis, Charles University in Prague and Academy of Sciences of the Czech Republic, May 2008.

H. von Stackelberg. *Marketform und Gleichgewicht.* Springer-Verlag, Berlin, 1934.

D. De Wolf and Y. Smeers. A stochastic version of a stackelberg-nash-cournot equilibrium model. *Management Science*, 43(2):190–197, 1997.

H. Xu. An mpcc approach for stochastic stackelberg-nash-cournot equilibrium. *Optimization*, 54(1):27–57, 2005.

H. Xu. An implicit programming approach for a class of stochastic mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 16(3): 670–696, 2006.

H. Xu and J.J. Ye. Necessary optimality conditions for two-stage stochastic programs with equilibrium constraints. Submitted, 2009.

J. J. Ye. Optimiality conditions for optimization problems with complementarity constraints. *SIAM Journal on Optimization*, 9:374–387, 1999.

J. J. Ye. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *Journal of Mathematical Analysis and Applications*, 307:350–369, 2005.

J. J. Ye and X.Y. Ye. Necessary optimality conditions for optimization problems with variational inequality constraints. *Mathematics of Operations Research*, 22:977–997, 1997.

R. Zhang. Problems of hierarchical optimization in finite dimensions. *SIAM Journal on Optimization*, 4:208–227, 1994.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 28.01.2008 Thomas Michael Surowiec